序

本书的合作者奥托·托普利茨，在1939年春离开德国后，于1940年2月19日在耶路撒冷逝世。托普利茨是在哥廷根时作为大卫·希尔伯特的弟子而开始其学术生涯的，其时他在基尔，随后又在波恩任教授。他的科研工作集中围绕在积分方程论和无限多元函数论方面，这是他长期从事并作出贡献的领域。

本书的梗概是当托普利茨在基尔，我在汉堡大学时，我们之间的频繁交往中产生的。有关本书的课题与内容，我们两人都向广大听众反复讲述过。我们相互切磋，多次改写了手稿。托普利茨对数学史有浓厚兴趣，这在现在的版本中仍然可以见到。我怀着愉快的感情，回忆起1929年夏，我们在波恩最后一起润饰了德文手稿的日子。对于现在这个英文版，我可以肯定托普利茨是会感到欣喜和自豪的。因为他常常提及希望能有这样一个英文版。

我要感谢译者赫伯特·朱克曼教授艰辛而卓越的工作，我不只赞赏他对德文版本的确切翻译，而且我还认为，通过他的译文，他引进了许多更符合英语读者习惯的论述。他还为本书增加了两篇（第15篇和第28篇），它们都忠实地反映了原书的写作精神。

汉斯·拉德梅彻

1956年，美国费城

导　言

数学，由于它的语言、记法以及看上去显得奇特的符号，就象一堵高墙，把它和周围世界隔开了。那座墙的背后在干什么，就其大部分来说，外行人是感到神秘的。他设想的是一些枯燥乏味的数字，是受铁的法则制约的了无生气的机械结构。另一方面，这座墙往往极大地限制了躲在墙里的人的视野。他惯于用特有的尺码来度量一切数学对象，并且还自鸣得意，以为没有什么亵渎的东西会进入他的领地。

折毁这座墙，用一般人也能欣赏的方式来介绍数学是可能的吗？难道不能使数学的欣赏扩大到那些“数学天才”的小圈子以外吗？如果数学天才只是指那些富有才华、有新的数学发现的人，那么确实只是极少数。同样的道理，在那些能作乐曲的人们中，只有极少数具有音乐天才。然而，懂音乐，甚至能仿制乐曲，或者至少能欣赏音乐的人，却是大量的。我们相信，能够理解简单的数学思想的人，相对来说，不会少于通常所谓的音乐爱好者，并且只要能去掉人们从幼年时代的经验中大量形成的对数学的成见，那么他们的兴趣就会大大提高。

这里各篇的目的在于表明，只要能抓住实质把基本的数学思想介绍清楚，对数学的成见就可消除。本书试图为包含

有数学的各种现象，为数学本身，并为其所具有的固有价值举出一些例子。

设法向不懂数学的人介绍数学是常有之事，不过通常的作法总是强调数学在人们所从事的其它领域中的用途，竭力确保读者的理解和兴趣。不厌其繁地叙述数学在技术和其它方面的好处，并用很多例子加以说明。另一方面，写出了许多有关数学游戏方面的书籍。虽然这些书搜集了不少有趣的材料，但它们恰恰严重地歪曲了数学的真实面貌。最后，另外一些书籍是按一般哲学的合理性来讨论某些数学的基础的。阅读以后各篇并对绝对的纯数学有初步爱好的读者，自然会把他的注意力放在这样一种数学的认识论的估价上去。但这好象是我们对数学附加了一种外来的价值，是根据它本身以外的标准来判断它的价值。

在以后各篇中，我们将不可能论证在数学本身范围内所提出的一些概念的意义。我们不能考虑数学内部的应用，就是说，把数学一个领域的概念和结果，应用于数学的其它领域。这意味着，我们必须略去数学体系中某些极本质的东西：渗入这个体系的各个方向的巨大而错综复杂的联系。这种省略在我们方面是极不愿意的，因为重大的数学发现往往就在于揭示这种十分重要的相互关系。但是，为了揭示这种相互关系，我们需要作长期的多方面的准备，而且对读者来说，还必须进行全面的训练。这些都不是我们这里打算作的。

换句话说，我们介绍的重点，不是那种其它学科能向一般人揭示的事实，而是包含有数学的现象的一些类型，提出问题的方法以及解决问题的方法。毫无疑问，若要理解重大的数学成果，是需要有渊博的理论，长期的练习和坚持不懈

的努力的。这对于音乐也是如此。第一次参加音乐会的人，难以欣赏巴赫的"赋格曲"，也不能立即具体了解交响乐的结构。但除了大型的音乐作品外，还有一些较小的然而确实是绝妙的片段，而它们所含的精神却能使人人心领神会。我们打算从广阔的数学领域中选取这样一些"较小的片段"：一系列的逐一都是完整的课题，每一题的阅读和了解不用超过一小时。这些题目都是独立的。因此阅读任一篇时，不需要记住以前读过了什么。也不要求读者去回忆他青年时期可能不得不学的那些东西。用不着对数和三角学。也没有提到微积分。几何上全等的理论，以及代数和的乘法将会逐渐被带回到读者的记忆中来；这就是一切。

至于一件小的音乐作品，它也许不只是依靠使它得以展开的乐谱线才显得美妙的。主旋律的一点细微变化，一个意外的变调都能成为全曲的高潮。在类似的意义下，我们的读者在基本思想出现明确变调以前，似应认真仔细"倾听"问题的基本目的，它的发展，以及用以说明每个主题的最初的几个例子。在进行推理时，他应当比通常在阅读时所要求的更积极主动些。如果能做到这些，他将发现，掌握每一个题目的基本思想并不困难。这时他会恍然大悟于一些伟大思想家所创作的东西。当时他们偶然地离开了他们深刻的理论成果的领域，从简单的开端出发，制作了一种小型独立的艺术作品，即短篇数学杰作。

目　　录

1. 素数序列

6 等于 2 乘 3，但 7 不能同样写作两个因子之积，所以 7 叫作素数（或称质数）。一个素数是一个正整数，它除了本身和 1 以外，不能写作两个较小的因子之积。5 和 3 也是素数，但 4 和 12 不是，因为 $4=2\cdot2$，而 $12=3\cdot4$。象 4 和 12 这种可以分解因子的数叫作合成数。数 1 不是合成数，但它又和其它数是如此地不同，因而通常也不把它看成素数；因此 2 是第一个素数，开头一些素数是：

2，3，5，7，11，13，17，19，23，29，31，37，……

一眼就可以看出，这个序列不符合任何一种简单规律，实际上，素数序列的构造是非常复杂的。

一个数可以逐步分解，直到它化为若干素数之积。象 $6=2\cdot3$ 立即表为两个素数之积，而 $30=5\cdot6$，又 $6=2\cdot3$，得出 $30=2\cdot3\cdot5$，为三个素数之积。同样，24 有四个素因子，$(24=3\cdot8=3\cdot2\cdot4=3\cdot2\cdot2\cdot2)$，其中恰好有 3 个因子是相同的素数 2。对于一个素数来说，只能写成 $5=5$，即一个单独素数之积。用逐步分解的方法，除 1 外的任何正整数都可以写作素数之积。因此，素数可以看作是全部正整数序列的基础。

在欧几里德的《几何原本》第九卷中，提出并解答了素数

序列最后是否有终结的问题。已经证明，这个序列没有终点，也就是在每一个素数之后，总能找到另一个更大的素数，

欧氏的证明非常巧妙而又十分简单。一些数 3，6，9，12，15，18……都是 3 的倍数。其它的数都不能被 3 整除。依次较大的数 4，7，10，13，16，19……都是 3 的倍数加 1，自然是不能被 3 整除的；例如，$19=6\cdot3+1$，$22=7\cdot3+1$ 等等。同样，5 的倍数加 1 是不能被 5 整除的（$21=4\cdot5+1$）。对于 7，11 等等也都一样。

欧氏写出了以下各数：

$$2\cdot3+1=7$$

$$2\cdot3\cdot5+1=31$$

$$2\cdot3\cdot5\cdot7+1=211$$

$$2\cdot3\cdot5\cdot7\cdot11+1=2311$$

$$2\cdot3\cdot5\cdot7\cdot11\cdot13+1=30,031, \text{等等,}$$

头两个素数，头三个素数，以及更多的几个素数乘在一起然后再加 1。这些数不能被任一个组成它的素数整除。因为 31 是 2 的倍数加 1，它不能被 2 整除。它又是 3 的倍数加 1，所以不能被 3 整除，它又是 5 的倍数加 1 所以也不能被 5 整除。31 恰好是素数，并且一定大于 5。211 和 2311 是新的素数，但 30031 不是素数。然而，30031 不能被 2，3，5，7，11 或 13 整除，所以它的素因子一定大于 13。事实上，通过简单计算，$30031=59\cdot509$，而这些素因子都大于 13。

这样的论证我们可以照样推演下去。令 P 为任一素数，作出由 2 到 P 的全部素数的乘积再加 1，写成

$$2\cdot3\cdot5\cdot7\cdot11\cdots P+1=N,$$

素数 $2,3,5,\cdots P$ 中没有一个可以整除 N，这样 N 或者是素

数（自然是大于P的），或者N的全部素因子都和$2,3,5,\cdots P$不同，并且大于P。不论在何种情形，一个大于P的素数已经找到。因此，不管P有多么大，总有更大的素数存在。

欧氏这部分论证很值得注意，而要说明它的绝妙特征却是有困难的。这个问题本身具有理论上的兴趣。而且只有对数学思想有一定内在感情的人，由于个人的爱好，才会提出这个问题。这种对数学的感情和对数学的美妙的欣赏，在古希腊时代是非常明显的，而且已传给了后代的文化。还有，这一问题是绝大多数人所轻视的。正当这一问题引起我们注意的时候，它显得仿佛是琐碎而多余的，而它的实际困难是不能立即看出来的。最后，我们必须赞扬欧氏证明这一定理所用的巧妙而简单的方法。试图证明这一定理的较自然的方法不是欧氏的方法。较自然的方法是试求任一已知素数后边紧跟的那个素数。但是由于素数组成的极端的无规则性，所作的这种尝试最后都失败了。

欧氏的证法克服了素数序列的组成没有规律的障碍，就是用某一个大得多的素数去代替P后面的下一个素数。例如，在他的证明中，过了素数11，得出2311，而不是素数13。并且过了素数13，得出素数59。在所考察的素数和证明中所得到的素数之间，常有很多的素数。这并没有削弱这个证法的意义，而只能证实希腊人当时的智慧还不能使他们作出更多的贡献。

作为素数序列复杂性的一个证据，我们要指出在序列中有着很大的间隙。例如，我们将证明，可以找到1000个相邻的数，它们都是合成数。这个方法是和欧氏法密切相关

的。

我们知道 $2\cdot3\cdot5+1=31$ 不能被 2，3 或 5 整除。如果两个数都能被 2 整除，则其和当然也能被 2 整除。同样，对于 3，5 等等也是如此。现在 $2\cdot3\cdot5$ 可以被 2，3 和 5 整除，因而 $2\cdot3\cdot5+2=32$ 可被 2 整除，$2\cdot3\cdot5+3=33$ 可被 3 整除，$2\cdot3\cdot5+4=34$ 可被 2 整除，$2\cdot3\cdot5+5=35$ 可被 5 整除，$2\cdot3\cdot5+6=36$ 可被 2 整除。所以，数 32，33，34，35，36 都不是素数。这个论证用于 $2\cdot3\cdot5+7=37$ 时第一次不能成立，它不能被 2，3 或 5 整除。

用相同的方法，我们可以找到 1000 个相邻合成数。令 P 为第一个四位素数(1009)，作出下列 1000 个数

$$2\cdot3\cdots P+2，2\cdot3\cdots P+3，\cdots\cdots，2\cdot3\cdots P+1001。$$

可见 2，3，4，5，……，1001 中每一个数可以被素数 2，3，……，P 中之一整除，而 $2\cdot3\cdots\cdots P$ 也是这样。因此，上表所列每一个数，可被 2，3，……P 中之一整除，因而它们都不是素数。我们就是这样找出了都不是素数的 1000 个相邻数。所以，在素数序列中至少可以找出 1000 个数的间隙。

自然，在沿着素数序列走得相当远以前，这样长的间隙不会出现，但是，如果我们走得足够远了，我们可用同样方法得出我们所需要的长间隙。

这个问题和欧氏的问题在性质和证法上都很相似，不过关于素数序列中的间隙问题，古希腊人可没有研究过。它是近代数学家在研究大量其它有关问题时得到的。这些问题中的大多数是不容易解决的，有些到现在仍未能解决，而有些则把我们引向数学的若干全新领域。

让我们来考虑这些深入的问题中的一个，它仍然可以用

同样方法掌握，而且对于其它某些问题也能得到一些启发。3 的倍数是 3，6，9，···，而这些数加 1，即4，7，10···，前面已经提到了。其它的 2，5，8，11，14，17，20，23，···，都是除以 3 后余数是 2 的数。上面这些数是否包含无限多个素数？就是说，是否序列 2，5，11，17，23，···也没有终结？我们将证明，这样的素数是无限的。

首先我们必须指出，如果在序列 1，4，7，10，13······中任取二个相乘，则其积必仍是这一序列中的数。这些数中的每一个都是某一整数乘以 3 再加 1。设第一个数是 $3x+1$，第二个数是 $3y+1$，则它们的积是

$$
\begin{aligned}
(3x+1)(3y+1) &= 9xy+3y+3x+1 \\
&= 3(3xy+y+x)+1, \qquad (1)
\end{aligned}
$$

它仍然是 3 的倍数加 1，所以仍然是在原序列之中。

现在，对于 2，5，8，11······中任何一个可分解为素数乘积的数，其素因子中至少有一个必然仍是 2，5，8，11，···中的一个。例如 $14=2\cdot7$ 有因子 2，又 $35=5\cdot7$ 有因子 5。为了表明这总是正确的，我们注意到，一个素因子或者在 3 的倍数这类数中，或在序列 1，4，7，10······中，或在序列 2，5，8，······中。3 的倍数中只有 3 本身是素数，而它不能整除我们选定的数。如果全部素因子都在 1，4，7，10······一类数中，则按照上面指出的(1)，我们的数仍将属于这一类。因而一个数不属于这一类时，至少有一个素因子必定是数2，5，8，11······中之一。

现在我们可以用欧氏证法进行，只是考虑用

$$2\cdot3\cdot5\cdot11\cdots P-1=M$$

去代替 $2\cdot3\cdot5\cdot7\cdot11\cdots P+1=N$。

M是3的倍数减1，就是说，是2，5，8，11……中之一数。同N一样，显然M不能被素数2，3，5，7，11，……P中的任一个整除。或者M是一素数（大于P的）或者其全部素因子都大于P。在后一种情形，至少有一个素因子是2，5，8……中之一，所以，在这两种情形下，我们都能找到属于此类数的大于P的素数。因此，序列2，5，8……中包含无限多个素数。

剩下的问题是序列1，4，7，10，13…是否也包含无限多个素数。很可能，2，5，8…包含无限多个素数，而1，4，7，…仅仅包含有限个素数。但事实上，这后一序列也包含无限多个素数，不过要证明这一点需要用全然不同的方法。在后面的一篇内，我们将得到关于这个方法的一些线索。之所以提到最后这些问题，是为了指出近代数学的问题和方法是如何同希腊时代的那些相连系的。不但在个别问题上是这样，而且在近代数学的整个领域内也是如此，在这些饶有兴趣的领域里，研究工作一直在进行着。

2. 曲线通行网

一家电车公司决定在不改变现有轨道情况下重新组织其线路系统。并希望轨道的每一段仅仅使用于一条线路；乘客可以由一条线路改乘另一条线路，最后到达他要到的任何地方。问题是：要使轨道的全部段落包括在内，并且每一段轨道只允许有一条线路通过，公司至少应设置多少条线路？

对于其轨道网如图1所示的一个很小的城市，问题很简

单，一条线路从A到B，第二条从C到D，两条都通过K。或者一条线路从A经过K到D，而第二条从B经过K到C，最后还可以一条线路从A经过K到C，而另一条从B经过K到D，每一种情况，都需要两条线路。包括4条线段AK、BK、CK、DK的一条环形线路是不可能的。当然，公司可以在R设一个新的转换点，即一条经过K的线路到R终止，从R到B作短程往返，成为一条新的线路。但这增加了所需线路的数目。增加更多的转换点，线路个数将随之增加。但我们对此不感兴趣，因为原来的问题是求尽可能少的线路数目。

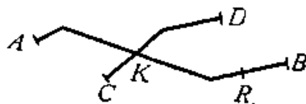


图 1

图2仍是一个十分简单的轨道网，但在这里问题也较为复杂。一种可能是，从A绕过B，C，D，E又回到A（一条封闭线路），第二条线路可以从A经过F，G，H到D,还需要的三条线路是BF，EG，CH，总共开出五条线路。但这并不是可能的最好安排。头两条线路可以合并成一条线路，即从A绕过B，C，D，E回到A后再经过F，G，H而到达D。这样，把线路数减少到四个。但是，是否可以只安排三条线路呢？

在图3的曲线网中，可以作从A绕过B，C，D和E回到A再经过F到B的一条线路，还有三条则是CF，DF，EF，而其中两条可合并成一条CFD，从而只剩下了一条EF。在这种情况下，头二条经过F，而第三条止于F。因而可以提出这样一个问题，能否最好安排两条线路呢？

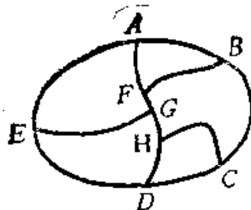


图 2

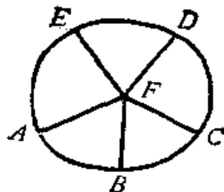


图 3

我们无需也很难把图 2 和图 3 中一切可能的线路都列出来（象对于图 1 那样），从而确定线路的最少条数。把复杂的曲线网列出来是一个很麻烦的程序，这在任何情况下，都不是一个使人感兴趣的问题。把任一已知曲线网的全部可能性都列出来，并不比两个七位数相乘更有数学意义。数学的精神是让我们找出问题的实质而用它解决问题。

问题的实质是十分简单的。我们必须考虑各条线路的端点在什么地方。在图 1 内，线段的自由端点如 A，B，C 和 D 必然是线路的端点，因为在这一情况下，肯定至少有四个端点，而每一条线路不能超过两个端点（自然，闭合线路没有端点），显然，在四个端点间至少有两条线路。这里，经过简单的思考，就得出了上述的从考虑一切可能情况中得到的同样结果。

在图 2 内，这里没有自由端点，但有象 A 那样有三条线段在这里会合的点。这样的地方至少是一条线路的端点，因为两条线路决不可能在同一线段上。在图 2 内，这里有八个这样的点，所以必然至少有线路的八个端点。八是一个偶数，即四条线路的端点。所以至少有四条线路。象我们在上面已提到的，正好是四条线路。

在图 3 中，有五个我们已讨论过的那样的连接点，此外还有第 6 点 F，这里不是三条而是五条线段在此会合了。我们称 F 这样的点为五阶的连接点。如在前面图 3 的讨论中所作的那样，在 F 处，两对线段可以连成线路，只剩下一段。同样，对于任意奇阶连接点，类似的讨论也能成立，即剩下了一个线段，因而这至少是一条线路的端点。现在我们看到图 3 中至少要有六个端点（是一个偶数），所以至少要有三条线路。

对于任一曲线网，不论多么复杂，都很容易数出奇阶连接点的个数，再除以 2，就得出可能的最少线路数。上述的三个例子中，奇阶连接点都是偶数个，这个数的一半就是线路数，这不但是问题的必要条件，而且是充分条件。现在我们要确定，是否对于任一曲线网，奇阶连接点的个数总是个偶数，是否这个数的一半就是线路数。

不管曲线网多么复杂，一定可以找出一些线路系统。在每一个这样的系统中，任一线段只能有一条线路通过。事实上，只要在每一对连接点之间的线段上都建立一条独立的往返线路，则就成为一个线路系统。不过，这就需要很多线路，因而还会有其它线路较少的系统。我们要找的是最优的，即所需的线路数是最少的。在一切可能设想到的系统中，必然有一个是最优的——即没有其它的系统比它需要更少的线路。

很明显，一个线路系统如果象图 1 那样，包括额外的转换点 R，当然不是最优的。又如果象图 3 中的 F 点那样，它是几条线路的端点，这也不是最优的。因为一条线路由 C 到 F，另一条由 D 到 F，则它们可在 F 连接成一条线路，从而

减少了线路的总数。为了使系统最优，连接点处的线段都要成对配成线路，这意味着在奇阶连接点处恰好有一线路在此终止，在偶阶点处就没有线路的端点。因而端点的总数，将是此曲线网中奇阶连接点的个数。

我们还将确定，一个最优系统是否可以包含闭合线路。这种闭合线路，我们在讨论图 2 时提到过，现在我们把它从 A 再连接到线路 $AFGHD$，使线路从五条减到四条。连接成的线路 $ABCDAFGHD$ 不是闭合的——它有两个端点 A 和 D。当某一闭合线路含奇阶连接点时，可以同样缩减线路数，而且对于仅包含偶阶连接点的闭合线路，也可作类似简化。设 A 是在闭合线路上的这样一个连接点（图 4）(这里是 8 字形)，经过 A 的其它线路以虚曲线表示，它们可以沿任意路径继续走下去。如果要使这一系统是最优的，则不能有线路终止于 A，所以由 B 经过 A 可以继续走下去，再经过 E，这样，我们可以把这条线路和闭合线路作成一个单一的线路，它通过 B 和 A，然后围绕闭合线路回到 A 又再通过 E。因为这样减少了线路的条数，所以原来的系统将不是最优的。不过，我们要注意，在这种简化中，新作成的线路可能是闭合的，因为原来的线路 BAE 可以继续下去，并且回到 B。如果是这样，进一步的简化仍是可能的。如果在某一步骤中，新的，包含一个奇阶连接点的闭合线路出现了，下一步的简化将产生一条非闭合的线路。不然的话，原曲线网

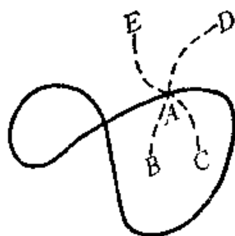


图 4

的全部连接点必然全是偶阶的，而这一系统最后将缩减为一条单一的闭合线路了。

总结上述结果，我们看到，如果一个系统是最优的，那么线路的端点只能是奇阶连接点，而每个这样的点只能是线路的端点。还有，最优系统一般不能有闭合线路，除非原曲线网的全部连接点都是偶阶的，而这时唯一的闭合线路将走遍整个曲线网。这样，在最优系统中，奇阶连接点的个数等于线路端点的个数，因此是一个偶数，而且线路的最小个数应是奇阶连接点的个数之半，除非全部连接点都是偶阶的，而这时的最少线路数，是一条（闭合的）线路。

3．一些极大问题

1．我们试比较周长为2吋的各种矩形的面积（图5）。每个矩形的宽都小于1吋，当宽接近于1吋时，则高和面积就变小。如果高接近于1吋，则其宽和面积也很小。高和宽都适中的矩形就有较大的面积，我们要问，究竟哪个矩形面积最大。这是一个关于极大的问题。大约是一切这类问题中最简单最古老的一个，因此用它作导引是最合适不过的了。

这个问题在欧氏著作卷Ⅵ的定理27中已解决了。我们的证明将应用同一原理，不同于欧氏的仅在于说明上。图6中矩形$ABCD$表示周长一定(P)的任一矩形。正方形$BEFG$每边长都是$\frac{1}{4}P$，所以它的周长也是P。我们断定这个正方



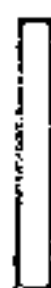


图 5

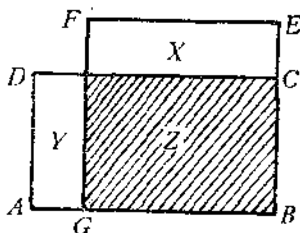


图 6

形就是问题的答案，就是说，它的面积比任何有相同周长的矩形的面积都大。在图上有阴影的矩形 Z 是原矩形和正方形共有的部分，正方形还包含一部分面积 X，而原矩形是由 Z 和 Y 组成的。现在 $AB+BC$ 是矩形周长之半，而 $GB+BE$ 是正方形周长之半，所以二者应相等；$AB+BC=GB+BE$。这可以写成 $AG+GB+BC=GB+BC+CE$，由此得到 $AG=CE$。这样，矩形 X 的高恰好是矩形 Y 的宽。但是 X 的宽是正方形的一边，而 Y 的高则是正方形边长的一部分，因



图 7

而比较小。如果图 7 中的二个矩形有一边长度相等，则另一边较大的矩形必有较大的面积，即 X 大于 Y。因此 $X+Z$ 大于 $Y+Z$，即正方形比矩形面积都大。只有当 $ABCD$ 成为正方形本身时，Y 的高不再是正方形边长的一部分，正方形 $BEFG$ 本身就是原始的矩形 $ABCD$ 了。所以**正方形面积比一切其它具有相同周长的矩形的面积都大**。

这个结果是古希腊人曾经叙述过的。它还可以用一个代数公式写出来，这样可以把它放人现代数学中。令 x 和 y 是用吋表示的矩形的二边长，矩形的面积可以写成平方吋数，而其周长为 $x+y+x+y=2(x+y)$ 吋。所以，正方形边长为 $\frac{1}{2}(x+y)$ 吋，而其面积为 $\left(\frac{x+y}{2}\right)^2$ 平方吋。这样，如果 x 和 y 是任意二个正数，其结果就成为：$xy \leqslant \left(\frac{x+y}{2}\right)^2$ 或 $\sqrt{xy} \leqslant \frac{x+y}{2}$[1]，用语言表达，可以这样说：两数的几何平均值常小于它们的算术平均值，只有两数相等时它们才相等。

2．现在我们已经看到所谓极大问题及其解法的意义了。首先提出了一个解，然后证明这个解所对应的图形，在已知性质上（这里是面积）与其它各图形比较都超过了它们。现在，可以转向这一节的主要问题：在内接于给定圆的三角形中，求面积为最大者。这个问题在欧氏以前一个世纪的柏拉图时代如果未能解决，也很可能已讨论过。然而，既不是欧几里德，也不是更近代的著作做出了下述的解答。它是由古希腊人研究和发现的。

在原三角形 ABC 之外，我们注意等边三角形 $A_0B_0C_0$ 内接于同一圆或另一等圆（图 8）。因为等边三角形除了可以转动外是稳固的，所以 $A_0B_0C_0$ 的面积已经完全确定。我们断定等边三角形是我们的解，就是说，它的面积比其它任何内接三角形都大。

我们首先注意到，等边三角形把已知圆周分成三个相等的弧，而其它三角形都分成不等的三个弧。后面的三个弧必

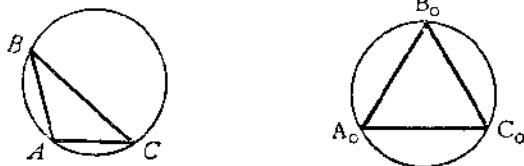


图 8

定有一个其长度大于圆周长的三分之一。若不是这样，为了组成全圆周，三个弧都将等于三分之一圆周，因而 ABC 将是等边三角形，但我们的假设并非如此。同样道理，至少有一个弧其长小于圆周的三分之一。第三个弧可能大于或小于圆周的三分之一；我们虽然不能肯定究竟是大还是小，但这并不影响我们的论证。

假设这个三角形已标明是这样的：弧 AB 小于圆周的三分之一，而弧 BC 大于三分之一。在图 9 内，沿着 CB 取弧 CB''，并使其长等于 AB。三角形 CAB'' 是 ACB 关于垂直于 AC 的直径的镜面反射像。我们再量弧 AB' 使其等于

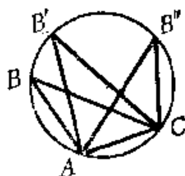


图 9

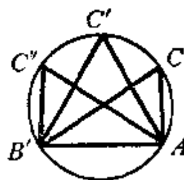


图 10

三分之一的圆周长，其方向和由 A 到 B 一样。因 AB 弧小于圆周的三分之一，B' 必然超过了 B。不过 B' 不能超过 B''，就是说，它落在 B 和 B'' 之间。如果它超过了 B''，则弧 AB'

将大于 AB''，而 AB'' 是 CB 的镜面反射像。但 CB 是大于三分之一圆周的，而我们已使 AB' 恰好等于圆周的三分之一。

今因 B' 在 B 和 B'' 之间，它比 B 高。三角形 ACB 和 ACB' 有同一底 AC，而 ACB 之高小于 ACB' 之高，因三角形面积等于高乘底的一半；所以 ACB 比 ACB' 的面积小。我们找到了一个新的内接三角形 ACB' 大于原来的三角形 ACB，而其一边 AB' 等于内接等边三角形的边长。

可能 ACB' 正好是一等边三角形。这将是在原三角形内截取的弧 AC 恰好等于圆周三分之一时的情形。这样，等边三角形大于原三角形的证明已经完毕。如果 ACB' 不是等边三角形，我们可把 AB' 作为图形之底，就象先前以 AC 为底一样。为此，我们转动图形，如图 10 所示，直到 AB' 作了底。我们现在考虑 $AB'C$，以 AB' 为底，就象对 ACB 以 AC 为底一样。我们最后得到等边的三角形 $AB'C'$，因为 $B'C'$ 和 AB' 都等于圆周的三分之一。由于图 8 中 $A_0B_0C_0$ 和 $AB'C'$ 都是内接于等圆的等边三角形，它们的面积是相等的。我们已经证明了 $AB'C'$ 大过 $AB'C$，而 $AB'C$ 又大过原三角形 ABC，而这个 ABC 假定不是等边的。这就完成了证明：在同一圆内，等边三角形面积大于任何其它内接三角形。

3．上面的结果不过是比较一般的论述的一个特殊情况，这一般论述可用同样方法证明出来。我们将证明，全部内接于给定圆的 n 边多边形中，正多边形是最大的。若 $n=3$，多边形化为三角形，就是上面的结果。

要证明这点，需要注意另外一个简单事实。如果已知内接于圆的任一多边形（图 11），我们可以画出另外一个内接

多边形，它有相同的边，但边的次序已任意安排过。我们只需画出到多边形角顶的半径，并沿着半径把圆分成几个扇形。这些扇形可按所需另排成其它次序。显然，新的多边形



图 11

与原多边形面积是相等的。

现在的证明与以前一样。首先，如果多边形是非等边的，则一定有的边割出的弧小于圆周的 n 分之一，也有的边割出的弧大于圆的 n 分之一。对于三角形来说，任何两边都是紧挨着的，但如 n 大于 3，我们感兴趣的两边可能就不是挨着的。不过，我们刚才看到，在同一圆内可作出一个有相同的边和面积的新的多边形，而它的两个特定的边是互相挨着的。假设小边叫作 AB，较大的是 BC。我们可以从 A 向 B 的方向量出圆周的 n 分之一，得一端点 B'。象先前 B' 应在 B 和 B 的镜象 B'' 之间一样，新的多边形将以 B' 代替 B，而其它顶点不变。这时它的面积就较原多边形加大了。同样，新多边形的一个边 AB' 是正多边形一边的长度。我们可以对其余 $n-1$ 个边采取同样步骤，在必要时可以重新排列边的次序。继续这种步骤，将得到一个正多边形，这是因为在每一步，新边的长度都正合适。因为每次面积都得到增

加，所以正多边形比原来的任意多边形的面积都大。

同样方法，可以证明所有外切于圆的多边形中，正多边形有最小的面积。

[注] 符号$<$的意义，以及读法都是“小于”；$\leqslant$的意义是“小于或等于”，或“不大于”。例如$3<5$，$3\leqslant 5$，并且，符号$>$的意义是“大于”，$\geqslant$的意义是“大于或等于”。

4. 不可通约线段或无理数

有关长度，面积和体积的度量是全部几何的基础。用一个线段去度量另外一个线段，我们看出不过是它的多少倍而已。量的时候如果恰好没有剩余，这就很简单。如果较小的线段不能把较大的线段恰好量完，则我们要看余量。可能余量恰好是度量线段的一半，三分之一，三分之二或其它分数。如果是这样，则我们改用另一种替代的度量线段，这是原来度量线段的一个分数部分。这个新线段是原来两个线段的“通约量”。

最早的几何问题无疑是包含通约量的。例如，一个矩形，其边长分别为3和4吋，按照毕达哥拉斯定理，在对角线上的正方形其面积为

$$3^2+4^2=9+16=25\text{ 平方吋。}$$

因此，对角线之长为5吋。因而一吋是矩形较小的边和对角线的通约量，它们的比为3:5。

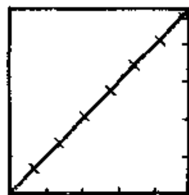


图 12

这样就很自然地把这一问题转到正方形去，看边和对角线的通约量。如果要想这样做，则需要求余量（图 12），从而迫使我们要试验越来越小的线段分数部分。这样产生了一个问题，即是否最终能找到一个充分小的通约量，或者竟永远没有这样一个通约量，换句话说，是否两个线段是“可通约的”或是“不可通约的”。这又提出一个问题，即是否一个线段可以分割为任意细微的小段，或者说是否有一个可能细分的极限。一条直线是由无数不可再分的小份所组成的吗？它是否有一个“原子结构”？物质的原子结构概念归功于德谟克列特，他出生在柏拉图之前。然而，这和直线的原子结构是有区别的。我们容易认为直线是“连续”的，就是说，可以分成任意微细的小份，但仍然假设直线是由一系列原子所串起来的片断。差不多同时保存下来的这个片断概念，应归功于阿那克萨哥拉；实际上，这是说直线是连续的。这个片断论流传下来的事实，与其说这不是一个偶然的说法，倒不如说，这是一个引起了很多讨论的有争议的论点，而且代表了人类在朝着解决这个基本问题的方向迈出了实际步子的那个时期。

我们可以想象，正方形的边和对角线是不可通约的这个影响深远的发现的巨大作用。这个发现应归功于毕达哥拉斯学派，这是南意大利的一个秘密协会，关于它所知甚少。据说研究这个问题的毕氏学派的一员在一次船只失事中死亡，从而作了公开赎罪。或许，这只不过是寓言，不是真的，因为对于当时的思想基础来说，发现无理数会引起极大的混乱。不管怎样，我们有柏拉图在他的《法规》一书中的自己的报告，谈到当他首次学习这一发现时，他是何等的激动。

我们将为这个事实给出两个证明，而不去考虑有趣的历史问题，即哪一个证明更早些。第二个证明不仅是由欧几里德，甚至是由亚历士多德提出的。第一个显然更古老一些，属于古希腊的典型，其精神见于欧氏著作的第十卷。

对于第一个证明，我们必须首先注意初等几何的一些事实。整个论证是容易理解的，即徒劳地企图找出一个公度量，先量正方形之边，再量对角线，不断作下去，试图找出一个合适的分数部分。我们由 B 沿对角线量出一边之长（图13）。只量一次，其端点止于 D。垂直于 BD 画线 $B'D$，且 B' 是与 AC 的交点。把 B 和 B' 也连结起来。我们有 $BA=$

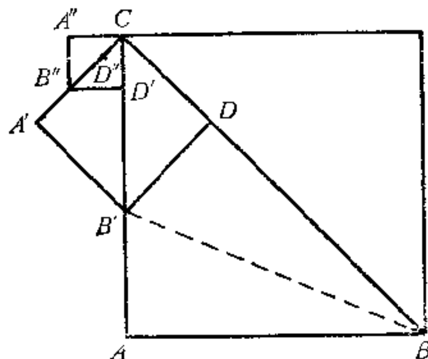


图 13

BD，$BB'=BB'$，$\angle BAB'=\angle BDB'$，最后一个结果是因为二者都是直角，所以由全等定理，三角形 BAB' 和 BDB' 是全等的，而 AB' 和 DB' 为对应边，因而相等。还有角 $B'CB$，在正方形一边和对角线之间，所以是半直角。又因角 CDB' 由作图看是一直角，在三角形 CDB' 内所剩的第三

角也是半直角，因此 CDB' 是一等腰三角形，其两边 DB' 和 DC 相等，总结以上的证明，我们有

$$AB'=B'D=DC。\tag{1}$$

现在我们由 C 点作对角线的垂线 $A'C$ 使其等于 DB'。A' 与 B' 相连结，作成 $A'B'DC$，是一个新的较小的正方形。对于这个新的正方形，以上全部步骤可以照样实行。在对角线上从 B' 量出一边之长得点 D'，再画 $B'C$ 的垂线 $B''D'$。如前我们有

$$A'B''=B''D'=D'C。\tag{2}$$

显然这是可以依次重复的，每次在对角线上有一剩余线段，而我们就拿它作为次一正方形的一个边。这个步骤是永远不能完结的，余段(它永远不能为零)在每步变得越来越小：

$$CD>CD'>CD''>CD'''\cdots\cdots,\tag{3}$$

每个余段是对应正方形对角线和边长的差：

$$CD=CB-AB,\ CD'=CB'-A'B',$$
$$CD''=CB''-A''B'',\cdots\tag{4}$$

这完成了第一证法的几何方面的准备，这个证明是反证法。我们假设边和对角线是可通约的，再证明这必将导致一个不可能的情况。如果二者是可通约的，它们必然有一个公度 E，它必然恰好把边和对角线都量尽。现在，如果任意两个线段都恰好是 E 的倍数，则其差必然也是 E 的倍数。所以，由(4)，如果 CB 和 AB 都是 E 的倍数，则 CD 必然也是。还有由(1)，我们有 $CB'=CA-B'A=AB-CD$，它是 E 的倍数之差，因而也恰是 E 的倍数。图 13 中第二个正方形有边 $A'B'=CD$ 和对角线 CB'，二者都是 E 的倍数。从第一个正方形到第二个，我们可以用同样步骤进行下去；由

$A'B'$ 和 CB' 可以找出 CD'，$A''B''$ 和 CB''，都是 E 的倍数，往后的正方形可以这样继续作下去。

我们现在遇到了矛盾。如果 CB 和 AB 是 E 的倍数，可以看出 CD'，CD''，CD'''……也都是 E 的倍数。但由(3)，这许多 E 的倍数越来越小而不能终止，即永不能为零。这是不可能的，因为例如，设 $CD=1000\,E$，则 CD' 有一个较小的 E 的倍数，最多是 $999\,E$，如此等等，直到最后的第 1001 项将小于 E。它将是零，因为它是小于 E 的，且是 E 的倍数。这就和(3)中没有一项是零的事实相矛盾了。

第二个证法就更简单了。它的算术方面的准备比第一证法中几何方面的准备简单多了。

我们先看偶数和奇数。一个偶数是另外一个数的二倍，所以可写作 $2x$。一个奇数是一个偶数加 1，所以可以写作 $2x+1$，一个奇数的平方永远是奇数，因

$$(2x+1)^2=4x^2+4x+1=2(2x^2+2x)+1,$$

它也是一个数的二倍加 1，由此我们可以立即证明：

引理 1.[1] 如果一个数的平方是偶数，则这个数本身必是偶数。因为如果这个数是奇数，象刚才看到的，它的平方也将是奇数。下面的引理更容易证明。

引理 2. 一个偶数的平方常可以被 4 整除。因 $(2x)^2=4x^2$ 是 x^2 的 4 倍，或简化为 $4g$。

主要的证法也是反证法。我们仍然假设，正方形的边和对角线有公度 E，则应用毕氏定理于正方形的两边和对角线组成的一个直角三角形，我们有

$$d^2=s^2+s^2,\quad d^2=2s^2. \tag{5}$$

假设 d 和 s 无公因子，因为选择新的 E，就必可把它们

化简。例如10和16有公因子2，但是把公度E二倍起来就可以把它们化简为5和8了。此后我们假设，这种简化都已完成。

由(5)，我们看到d^2是一个数的二倍，所以它是偶数。则由引理1，d也是偶数。因此s必为奇数，否则d和s将有公因数2，与已经把它们简化了的假设相矛盾，但因为d是偶数。由引理2证出d^2可被4整除，$d^2=4g$又$2s^2=4g$或$s^2=2g$，所以s^2是偶数从而s也是偶数(引理1的另一应用)。这与我们刚证明的结果(s为奇数)相矛盾。这就证明了原假设，即s和d有公度是错误的。

两个证法的实质是对于正整数递减序列最后必有终结。在第一个证法中，显然我们就是研究序列(3)，在第二个证法中，它实际上隐含在有关数d和s化简的论述之中。两个数的简化形式的证明，可以通过一系列的递减步骤求得。

用现代的数学记法，公式(5)可以写作

$$\left(\frac{d}{s}\right)^2=2,$$

同样，最后结果可以写成为：没有分数(即没有有理数)$x=\frac{d}{s}$其平方是2。这也可以说成是：没有有理数能等于$\sqrt{2}$，或者说，$\sqrt{2}$是一"无理数"。

〔注〕引理是其重要性不足以叫做定理的命题，但在后面的证明中将要用到。

5. 垂足三角形的一个极小性质

我们将再考虑象在第三篇中所研究的那类问题，但这时

应更恰当地叫作极小问题。它将引入十分完善而简明的数学方法。这里的定理和证明是施瓦茨作出的。虽然这个定理是比较小的数学问题，但在这比较普通而极有意义的工作中，同样显示了这位大数学家的智慧。

1. 在考虑主要定理之前，让我们先研究一下光学方面的反射定律这一极简单的问题。众所周知，如果从 A 发射一条光线(图 14)投射到镜面 g 上，它将反射到 B，走的是使投射角等于反射角的路径。我们所要证明的是，光线所选择的道路 ADB 是由 A 触到镜面 g 到 B 的一切可能路径中最短的。这就是汽船要从 A 地到另一 B 地而要先到对岸 g 时所走的路径。我们且不谈为什么没有推理能力的光线所走的和领航员在适当思考之后选出的路径是相同的问题。我们要证明的全是纯数学的事实，即投射角和反射角相等的路径 ADB，要比其他任何路径 ACB 为短。

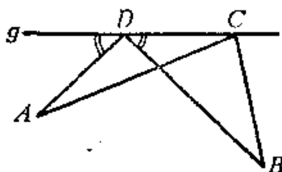


图 14

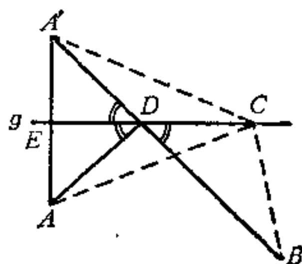


图 15

证明所依据的方法，从数学观点看仿佛是不自然的，但从光学观点看则是十分自然的。我们作 A 点及直线 AC 和 AD 关于镜面 g 的反射(图 15)。如果 A' 是 A 的影像，则 $A'C$ 是 AC 的影像，又 $A'D$ 是 AD 的影像，所以 $A'C=AC$，

$A'D=AD$ 又 $A'E=AE$，所以三角形 EDA 与 EDA' 是全等的，而角 EDA 和 EDA' 相等。按照我们的假设角 $EDA=$ 角 CDB。所以角 CDB 和 EDA' 应是对顶角，即 $A'DB$ 是一直线。

现在路径 ADB 和 $A'DB$ 以及 ACB 和 $A'CB$ 的长度都相等，因为 $A'DB$ 是连结 A' 同 B 的一条直线，它比路径 $A'CB$ 为短，从而 ADB 短于 ACB。这里我们应用了两点间直线最短这一事实。

2．现在我们转到主要问题上来，在一个已知锐角三角形 ABC 内，求作一个内接三角形 UVW，使其周界要尽可能地小(图 16)。结论是，垂足三角形 EFG (图 17)，即其顶点是三角形 ABC 三个高的垂足的三角形，与任意其它内接三角形 UVW 比较，它的周界是最短的。

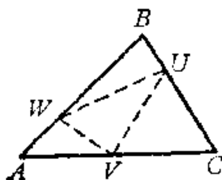


图 16

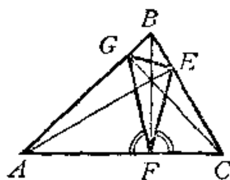


图 17

我们首先要证明关于垂足三角形的一个引理。角 AFG 和角 CFE 是相等的(象在反射律中一样)。在 E 和 G 处的类似的角也都相等。为要证明这个引理，我们要回忆平面几何中的一些定理。塞勒定理是：半圆上的圆周角是一直角(图 18)；同一弧所张的圆周角相等(图 19)；三角形的三个高相交于一点。利用这些，以 AH 为直径的圆通过 G 和 F，而以 CH

为直径的圆通过E和F(图 20)。而且，角AFG所对的弧为AG，与角AHG对的弧一样。所以这两角相等。同样方法，

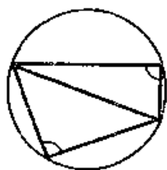


图 18

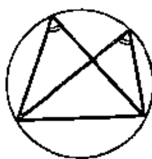


图 19

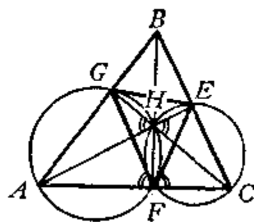


图 20

我们看到角CFE和角CHE也相等，但是角AHG和角CHE是对顶角，应相等。所以我们得出角AFG=角CFE。

3．现在我们可以开始谈施瓦茨的证明了。我们关于边BC反射三角形ABC(图 21)，再关于边CA'反射那反射过的三角形，再关于边$A'B'$反射那刚导出的三角形，如此关于$B'C'$，关于$C'A''$等，共作出六个反射。我们首先证明一个很明显的事实，即最终位置$A''B''C''$是从原位置ABC没有转动地平行移动后得到的。头两个反射把ABC移动到第三个位置$A'B'C$。这个移动实际上等于未作反射，也没有使三角形离开它所在平面，只是绕C点按顺时针方向转了一个角度$2C$，同样，由第三个位置移动到第五个位置就是绕B'按顺时针转过角度$2B$。最后，绕A''按顺时针再转过角度$2A$到达第七个位置，即最后位置。总计，三角形转了一周，即转过角度$2C+2B+2A$，因为$A+B+C$之和为三角形三内角之和，为一平角。所以最后三角形的位置与原三角形

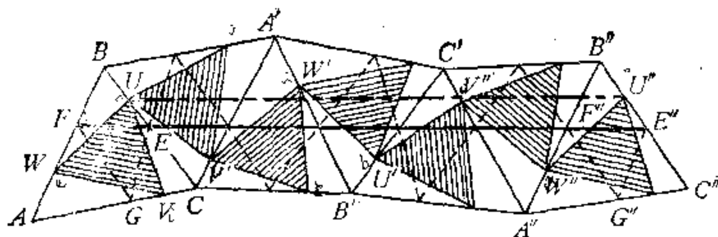


图 21

相同，只是作了一个平行移动，因而 BC 平行于 $B''C''$。

现在，我们想摘清楚在逐次反射之下，垂足三角形和三角形 UVW 的各个位置的踪迹。这在图 21 内用点线和阴影表示出来。从关于垂足三角形的引理，我们即可看到 EG 的第二个位置和 FE 的第一个位置形成一条直线；同理，垂足三角形的一个边的各次位置，将总是在这条直线上。所以直线 EE'' 是由六个线段组成的，2 个等于 FG，2 个等于 GE，2 个等于 EF：因而等于垂足三角形周长的二倍。

对于任意三角形 UVW，用同样方法追踪它的位置，我们得出一条锯齿线 $UV'W'U'V''W''U''$ 连结 U 和 U''，它等于三角形 UVW 周长的二倍。

我们可以看出线段 UE 和 $U''E''$ 平行，因它们在 BC 和 $B''C''$ 上，又因它们是三角形 ABC 两个位置的对应线段，所以也相等。由平面几何的定理 $EE''UU''$ 是一个平行四边形，因而其他二边相等，$UU''=EE''$。所以 UU'' 也是垂足三角形周长的两倍。UU'' 是连结 U 和 U'' 两点的直线，较连结同样两点的锯齿形线为短，而锯齿形线是 UVW 周长的二倍。所以，垂足三角形的周长较 UVW 的周长为短。

这个证法是一种典型的数学证明方法。其实质是在假设和结论之间作许多变换，直到定理的真正核心可以一望而知。

6. 前篇极小性的第二个证法

1. 在前一篇里，证明了内接于一个已知锐角三角形的一切三角形中，垂足三角形有最短周界。对于这一定理，还需要考虑另一个证法，因为这第二证法将说明一些新的概念。对于我们的目的来说，所用的方法将比一些新定理的数学内容更为重要而有趣。由施瓦茨最先提出的前一证法，主要根据的是两点之间的直线距离最短的事实，并应用了图形关于一直线反射概念。这两个原理也是第二证法的基础。对比它们在这二个证明中的运用方式，也是有趣的。下面的证法是费瑞提出的。他在学生时代发现了它，因而特别受到施瓦茨的重视。

2. 在已知锐角三角形 ABC 内(图 22)，令 UVW 为任一内接三角形，U 在 BC 上，V 在 CA 上，W 在 AB 上。令 U 分别关于二直线 AC 和 AB 作反射，而 U' 和 U'' 是其影像，因 UV 和 $U'V$ 互为镜面影像，因而相等。同理 UW 和 $U''W$ 也相等。三角形 UVW 的周长是 $UV+VW+WU$，因而它也等于路线 $U'V\ WU''$ 之长。

如果把 U 固定，但移动 V 和 W 到新位置，则点 U' 和 U'' 保持恒定，因为它们只和 U 和三角形 ABC 有关。连接两定点 U' 和 U'' 的路径 $U'VWU''$ 其长总等于 UVW 的周长。由 U' 到 U'' 的最短路径是一直线。所以，直线段

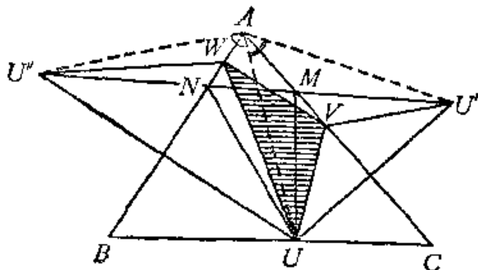


图 22

UU'' 是固定一个角顶U的所有内接三角形中可能的最小周长。这个具有顶点U的极小三角形，我们叫它为 UMN，如图 22 所示。

3．已经找到顶点在U的周长最小的三角形，我们只需要从U点在各个位置的三角形中选出其周长极小者。这个三角形自然是一切内接三角形中周长极短者。

我们必须确定U的这样的位置，以使线段 $U'U''$ 可能是最小的。为此目的，我们首先注意到三角形 $AU'U''$ 是以 AU' 和 AU'' 为等边的等腰三角形。事实上这两个线段，每个都是线段 AU 的镜面影像，所以相等，即 $AU=AU'=AU''$。

虽然三角形 $AU'U''$ 的两边都等于 AU 之长，而且和 U 在 BC 上的位置有关，角 $U''AU'$ 的大小却和 U 的位置无关。这个角完全由原三角形 ABC 所确定，因为(由于反射)在图形中诸角间有如下关系：

$$\angle UAB=\angle U''AB,\ \angle UAC=\angle U'AC.$$

从第一式，有 $\qquad \angle U''AU=2\angle UAB,$

从第二式，　　　　$\angle U'AU=2\angle UAC$；

所以　$\angle U'AU+\angle U''AU=2\angle UAB+2\angle UAC$，

或　　　　　　$\angle U'AU''=2\angle BAC$，

这证明了关于角 $U'AU''$ 的论断。

4．在等腰三角形 $U'AU''$ 内，我们将使底 $U'U''$ 尽可能地小。因为在 A 点的角和 U 无关，对于不同位置的 U，所有这些三角形有同一顶角。在所有这些三角形中，有最短的底者也必有最短的腰。腰 AU' 和 AU'' 其长都是 AU。所以如果取 U 使 AU 最短，则得到最短的线段 $U'U''$。

现在 AU 是连接点 A 和直线 BC 上一点的线段，我们知道，由一点到一直线垂直距离是最短的。所以我们要这样取 U，使得 AU 垂直于 BC，就是说，AU 是三角形 ABC 上通过 A 的高线。

5．现在让我们作出三角形 EFG 使其周界最短(图 23)。令 E 为过 A 的垂线的垂足。在关于 AC 和 AB 的反射之下，

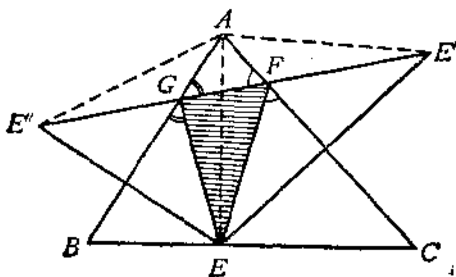


图 23

设 E' 和 E'' 是 E 的影像，则 $E'E''$ 是内接三角形最短周界之长，直线 $E'E''$ 与 AC 和 AB 相交于 F 和 G，这两个就是最

小周界三角形的其他两个顶点。

回想一下前面所提到的，我们看到，凡异于 EFG 的每一内接三角形 UVW，其周界必然是较大的。因为如果 U 异于 E，则线段 $U'U''$ 必大于 $E'E''$，而 UVW 的周界至少象 $U'U''$ 那样大。如果 U 和 E 重合，则点 V 和 W 中之一或二，必异于 F 和 G，而路线 $E'VWE''$ 将异于直线 $E'FGE''$。因此，在两种情况下，UVW 的周界必然大于 EFG 的周界。

6．这些讨论，证明了最短周界的内接三角形问题只有一个解。我们将应用这个解的唯一性。我们对于最小三角形的作图，不必对三个角顶同样去作。一个角顶 E 是过 A 的高的垂足，而在作图中，对其它两点不必再考虑由 B 和 C 作高了。

我们可以用 B 替代 A 作为起始点，并完成全部论证。就是说，如在第 2 节中一样，用点 V 关于边 BA 和 BC 的反射，代替点 U 关于 AB 和 AC 的反射，如此类推。我们最后将得到，过 B 的高的垂足下也有一个最小三角形。因为这里只有一个最小三角形，通过 B 点的作图必然也导出同一三角形 EFG，这同从 A 开始的原作图一样。我们也可以从 C 点开始，这样总结出最小三角形 EFG，不但 E，而且 F 或 G 都是高的垂足。这样，定理已经得证。

7．我们从证明中可以再多了解一点东西。应用解的唯一性，则如果 E 有某种性质(它是高的垂足)，那么 F 和 G 必也有类似的性质。同样，在作图中 F 和 G 的性质必然在 E 中也能成立。因为 E' 是 E 的镜像，角 EFC 和角 $E'FC$ 相等。因为角 $E'FC$ 和角 GFA 是对顶角，它们相等，又角 EFC ＝角 GFA。就是说，最小三角形的两个边在 F 处同原三角

形的边 AC 作出等角。对于点 G 相应的说法也能成立。如果从通过 B 点的高的垂足 F 开始我们的作图，同样证明角 GEB 和角 FEC 相等。

且不论三角形 EFG 的最小性，从第 6 节我们知道 EFG 的特征是一垂足三角形。结合这两个结果，我们得到这一定理：内接于锐角三角形的垂足三角形，在每一顶点的两个边和原三角形的一个边作出相等的角。

刚才这个定理，并无关于极小性的任何内容。它是初等几何学中常见的一个定理，它可以用初等几何方法证明出来。事实上在第五篇内，我们已经作过了。施瓦茨的证法需要这个结果作为引理，我们可利用它去求证关于圆的大量定理。费瑞的证法的一个优点是仅仅使用了最短距离和反射原理。而且费瑞的证明突出地只用了两次反射而施瓦茨的则用了六次。

8．关于垂足三角形的定理，还有一个类似的定理：在任意一个锐角三角形内必有一个，也只有一个这样的点，从这个点到三个顶点距离之和为极小，又从这点到三个顶点所连的三条线，彼此都作出 120° 的角。

这个定理是由舒拉特卡证明的，他提出的方法和施瓦茨对于垂足三角形定理的证法是类似的，一个更简短的证法是巴克纳提出的，我们将采用这个方法。

令 P（图 24 a）为锐角三角形 ABC 内任一点，令三角形 ACP 绕 A 点旋转 60° 到 $AC'P'$ 的位置，这个旋转是这样一个方向，使 AC 转到三角形之外，最后使得线 AC 在 AB 和 AC' 之间，这样我们有 $C'P'=CP$ 又 $PP'=AP$，对于三角形 APP' 不但是等腰的，也是等边的，它在 A 处的角是 60°。

所以路径 $BPP'C'$ 是 P 点到三个顶点距离之和，且点 C' 和 P 的位置无关。对应于 P 的不同位置的所有路径，其最短路径是直线 BC'(图 24 b)，所以最小点 P_0 必在 BC' 上，而其

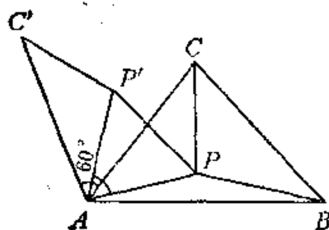


图 24 a

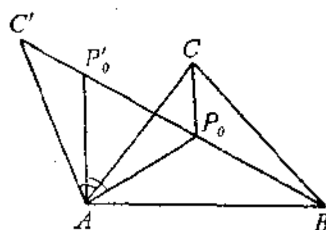


图 24 b

位置完全可由角 $AP_0C'=60°$ 的事实确定。其补角 AP_0B 因而是 120°。这个作图说明只能有一个极小点 P_0，用另一顶点代替 A，同样作图，将导出同一点 P_0，所以角 BP_0C 和角 CP_0A 也都是 120°。

7. 集 合 论

本篇讨论的集合论是数学的真正基础。然而，我们的兴趣，比起它的数学意义来，更多的是在建立它的方式的美妙和简明性上。集合论起源于康托，是从最简单的概念出发，利用纯粹的推理而建立起来的重要的数学分支，是真正的数学理论。

是否整数比偶数更多些？在一个线段上的点和在正方形面上的点，哪个更多些？康托在他的理论中，正是首先提出

了这样的问题。当试图回答这些问题时，重要的是要避免马上就提到结论。他所提出的问题不是很明确的，因为我们不确切知道一个集合多于另一个集合的意义。康托的第一个重要步骤，就是象对有限数那样，用简单的计数方法，给出一个明确的意义，并且细心地区别基数和序数，而对于有限数，这种区分仅是语法而已。

一个简单的例子将指出我们所要走的方向。假设有人在舞厅内问我们，这里是男人多还是女人多？怎样才是最容易的判断方法呢？一个方法是把男人和女人分成两组，对每组进行计数，再比较一下人数。但是，较简单的方法，将是开始跳舞，男人和女人将配起对来，只需要观察一下是男人还是女人剩下就行了。这里我们假设每人都参加了跳舞，如果他能找到舞伴的话。

康托的出发点，就是采用了配对的原则。如果想确定是整数多还是偶数多，我们不妨把它们配起对来。事实上，我们发现，每一整数和每一偶数配起来后并无剩余。我们得出如下：

$$\begin{array}{cccccc} 1, & 2, & 3, & 4, & 5, & 6, \cdots\cdots \\ 2, & 4, & 6, & 8, & 10, & 12, \cdots\cdots \end{array}$$

上行每一个整数都和下行直接相对的偶数配起来了。用这个方法，每个数都得到配对而没有剩余。这一简单事实很值得注意。两行的确匹配起来了，但下行仅仅是上行的一部分。

这引出了这种情况和有限数情况的一个本质区别。在舞厅的情况下(这里有有限多个舞客)，哪个男人和哪个女人伴舞是无关重要的。只要没有人进入和退出舞厅。不参加跳舞的人数总是一样的，这和整数和偶数的情况完全不同。我们

见到，使每个元素都得到配对的一种方法确实存在，但也容易得出另外一种配对法。最自然的配对法是：2对2，4对4，6对6等等，这样，把所有偶数都配起对来了，但留下了全部奇数。康托理论的实质是它放弃任意配对的想法，而只要求有一个使所有元素都恰好配对的方法。如果这样的配对法可以找到，说明这两个集是同势的。

康托的下一步就是证明，由全部整数和分数组成的有理数集没有比整数集更高的势。为了制造适当的配对方式，我们排列分数不是以其大小为序，而是依其分子分母之和的大小为序。我们只考虑化简过的分数，而以分子分母之和为2时开始，这只有一个分数：$\frac{1}{1}=1$。其和为3的是$\frac{1}{2}$及$\frac{2}{1}=2$。其次，和为4的是$\frac{1}{3}$，$\frac{2}{2}$，$\frac{3}{1}$，但$\frac{2}{2}$应取消，因它不是化简过的。此后，跟着有$\frac{1}{4}$，$\frac{2}{3}$，$\frac{3}{2}$，$\frac{4}{1}$，等等。我们按这个次序把分数写在自然数序之下，

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, \cdots\cdots$$

$$1, \frac{1}{2}, 2, \frac{1}{3}, 3, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, 4, \frac{1}{5}, 5, \frac{1}{6}, \cdots\cdots$$

把每个整数和直接在它下边的数配起对来。下面这一行不会遗漏任何有理数，因为每一有理数其分子分母之和必然是一有限数，而且必在此行之内。这就是我们所需要的配对法。

这个使人惊奇的结果表明，有理数集是"可数的"。因为它们同自然数配起来了，就是说，可以数出有理数来。一般，当一个集可与自然数集配对时，就是可数集，而且同自然数集有相同的势。康托继续证明，其它一些比自然数更广

泛的集，实际上不能有更高的势。我们将不再谈这些例子了，因为它们包含更多的数学概念，而且这一个例子已经够了。

这样，所有已经考虑过的无限集都有同一的势。如果没有更高的势的集，康托的理论就无足轻重了。康托证明，更高的势的集确实存在。这只要证明一线段上的点的集合比自然数集有更高的势。这里的证明是反证法。假设在一英寸线段上的点和自然数之间有一个配对关系，然后引出一个矛盾来。这一配对法可以使线段上的点采取这种次序，第一个配作1，第二个配作2，等等，但这显然不是线上的点的自然次序。把一个点和它同线段一个端点的距离看成一样将是方便的。例如，线段的中点相当于数0.5，而每一点相当于某个小数。为了量点的位置，我们必须用无限小数；例如，距离端点为线段三分之一的点相当于0.3333···。在线段上的点和自然数配对时，我们可用对应的无限小数以代替点本身。在第一个位置将有某个小数0.······对应于1，在第二个位置有类似的一个，等等。为了方便，可把配对关系垂直列表如下，其中列入的粗体数字，目的是为了说明。现在我们要指出，将有一个小数0.···不包含在此表内，这就和我们

1.	0.**3** 5 4 2 0
2.	0.6 **1** 7 7 3
3.	0.5 5 **5** 4 9
4.	0.0 1 0 **0** 7
5.	0.2 0 2 0 **6**
⋮	⋮ ⋮ ⋮ ⋮ ⋮ ⋮

的假设矛盾了。这个小数可以这样找到，我们选择小数点后第一个数字，不同于表上第一个小数的第一个数字。这样我们可在1到9这9个数字中选取。为确定起见，我们取数字1，而若第一个小数的第一数为1时，则取数字2。现在显然这个新小数不同于表中第一小数，而不管其它数字是怎样取的。同样，小数点后第2个数字取作1，除非表中第二个小数的第二个数字是1，而这时就取作2。所以，在这两种情况下，表中第二小数也不同于新小数。现在，新小数和表中第1及第2个小数都不同。用这个方法继续选取数字。在上例中将得到新小数0.12111···。因为这种步骤可以无限地进行，我们将定出一个小数来，它和表中全部小数都不同。这就说明表中不包含这一新作的小数，而这与它能得到配对的假设矛盾。所以，线段上的点是不可数的。

上一证法有需要加以说明的某种缺点。由于存在着尾部是无数个9的小数，如象0.269999···，这就与0.270000=0.27没有区别了。于是发生了两个不同的小数表示同一点的麻烦，从而表明我们不能用全部小数的集去代替线段上的点集。这两个集不能恰好对应。康托用一个形式十分不同的证明方法，避免了这一困难。但也有一个非常简单的方法去弥补这一缺点。这只要排除一切以9结尾的小数就可以了。即我们不用0.26999···去代表一个点，而总是用0.270000代表该点。我们所做的这一切，就是要保证所作的小数不能以9结尾。但对于证明中所作的新小数来说，这是不可能出现的，因为所作的小数全部是数字1和2所构成，并不包含任何9。

这个证明还有一个有趣的结果。因为有理数集是可数

的，我们并已知道它的势比 0 与 1 之间全部数的集的势低，因此，在 0 与 1 之间的数一定有不是有理数的数即必有无理数。所以，这就是用了十分一般的研究方法，证明了无理数的存在性。同样的事实在第四章中用完全不同的方法也已经证明过了。

康托的下一个结果也是令人惊异的。它说明一个正方形面上的点集，不能比正方形一边上的点集有更高的势。这个结果之所以奇异，其原因是它同维的直观概念相矛盾了。一维的线段和二维的正方形有相同的势；也可同样证明三维的正方体也有相同的势。

如同在前面的证法中一样，我们将采用无限小数，但仍然排除了尾部为 9 的小数。线段上的点用小数 0.……表示，和前面一样。正方形上的点将用一对小数表示，x 表示点与正方形左边的水平距离，y 则表示点与正方形下边的铅直距离(图 25)。我们将用下面方法来配对。正方形上每一点 P 将用两个小数

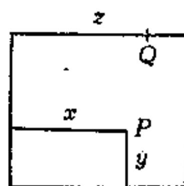


图 25

$$x=0.a_1a_2a_3\cdots,\quad y=0.b_1b_2b_3\cdots$$

表示。由这两个小数作出一个简单小数，

$$z=0.a_1b_1a_2b_2a_3b_3\cdots$$

即轮换取 x 和 y 的数字。今将正方形内的 P 点和线段上的 Q 配起对来，而点 Q 用 Z 表示。例如正方形中心，$x=0.500\cdots$，$y=0.500\cdots$ 和线段上相对应的点 $z=0.550000$ 配对，用这种方法正方形上每一点都将同线段上某个点配对。但只说配对还不够确切。如果我们仅仅说把正方形内每

一点，和线段上某点相配，我们也可以把正方形内每一点 P 和 P 点正上方的正方形上边线之点相配，但这样边上的每一点将不是恰好和正方形内一点相配，而是有无数多个点了(即铅直线上所有的点)。但在康托比较集的势的方法中把这一情况排除了。正象舞客的情况一样，不能假设一个人能有好多个舞伴。我们刚才那个比较严密的配对方法已经避免了这一困难。如果在正方形内有第二点 P'，其中

$$x'=0.a_1'a_2'a_3'\cdots,\quad y'=0.b_1'b_2'b_3'\cdots$$

它同线段上同一点 Q 相配，则我们有

$$z=0.a_1'b_1'a_2'b_2'a_3'b_3'\cdots$$

这个 z 的小数表示式与原表示式，只有在全部对应数字都相等时才相等，

$$a_1'=a_1, b_1'=b_1, a_2'=a_2, b_2'=b_2, \cdots,$$

但这表明 x'的全部数字等于 x 的对应数字。对于 y'和 y 也是这样。所以我们有$x'=x, y'=y$ 即 P'和 P 是同一点，这与它们不是同一点的假设相矛盾。因此，正方形上两个不同之点绝不能同线段上同一点配对。

当我们证明在线段上没有留下一个点没被配对时，我们的证明就完成了。这是容易看出来的，因为，如果

$$z=0.c_1c_2c_3c_4c_5c_6\cdots$$

对应于线段上任一点，则正方形内一点

$$x=0.c_1c_3c_5\cdots,\quad y=0.c_2c_4c_6\cdots$$

将与线段上某点配对，其对应的小数将是从 x 和 y 中轮换取出的，而这正是前面提出的小数 z。

在证明中仍然有一个缺点，仍来自尾部为 9 的小数。线段上一点 $z=0.2202020\cdots$与正方形中一点配对，有

$$x=0.2000\cdots,\ y=0.2222\cdots。$$

线段上一点 $z'=0.12929292\cdots$ 和正方形中一点

$$x'=0.1999\cdots,\ y'=0.2222\cdots$$

相配，但这里有尾部为9的小数，我们把它写成 $x'=0.2000\cdots$ 所以在线段上对应于 z 和 z' 的相异两点，和正方形内同一点相配了。

一个简单而并非一目了然的方法可以消除这一困难。上面所提到的那种形式，使得证明不正确了，但作如下修正后，证明就是正确的了。当9在小数中出现时，我们把它和它右边的数字合起来作成一个不可分割的“分子”。如果有很多相邻的9，我们把它们和下一个数字合成一个分子。这样，在前面的例子中，我们将有0.(1)(2)(92)(92)(92)⋯。举一个新例子，我们将写出 $z''=0.(7)(3)(94)(990)(9997)\cdots$，在证明中，我们以分子去代替数字，例如，由 z'' 我们有

$$x''=0.(7)(94)(9997)\cdots,\ y''=0.(3)(990)\cdots$$

现在的配对法和以前的有所不同了，但略一回想，由于除去了尾部为9的小数，我们看出在证明上不再有困难了。

在这里，康托遇到一个后来变得很出名的问题。这个问题就是：是否存在一个集，它比自然数集有较高的势而比线段上的点集有较低的势。

这个问题曾被称作“连续统问题”。它是对康托及其后继者寻求答案的一切努力的挑战。大概没有任何其它的数学问题，提出它只须具有很少的预备知识，但解决它却是如此困难。提出它所要用的知识不过是整数和线段的概念。如果要用一些复杂的数学概念才能提出某种很困难的问题，那是不足为奇的。仅仅用一些简单概念就提出了一个问题，它既不

容易解决，又出色地显示了数学的真正艺术。从这一观点考虑，连续统问题是一个光辉的范例。

不久就显示出来，集合的基本理论是全部数学的依据。但在集的概念未被仔细地分析以前，集论是不可能真正地发展的。由于集论中的一个著名的悖论，才迫使人们去注意这样的分析。

我们已讲到集，这些集是由这类或那类“元素”组成的。线段的点集，包含作为元素的线段上的各个点。整数本身是整数集的元素。集和元素之间的关系，就如同团体和它的成员一样。有时一个组织不以个别东西为其元素，而是以另一些组织为其元素。例如，联合国是各个国家的组织，而每一个国家又是公民的组织。联合国确切的成员是各个国家，而不是国家的公民。同样，一个集可包含一些集作为它的元素。例如，全部可数集的集，其所有的元素本身都是集。一个个别国家的公民并不是联合国的一元。类似，数$\frac{1}{5}$是有理数集的一元，但它不能作为全部可数集组成之集的一元。

一个集能否包含其自身作为一元？我们已经看见过的普通的集没有这个性质。然而，容易给出例子，说明这种非普通集是存在的。全部可设想到的集的集是这一类型的，因为它是包含自身的一个特殊集。这时，我们将把包含集本身作为一元的集叫作非普通集，而其它集叫作普通集。

现在让我们考虑全部普通集的集，而把它记作s。s本身是一普通集还是非普通集？它必然是其中之一。如果s是非普通集，则它一定包含其自身作为一元。但这时s象s的全部元一样，是s的一个元，因而它是一个普通集。这是一个

矛盾，因此 s 不是非普通集。然而，如果 s 是普通集，则它不能包含它自身作为一元。因而 s 不是 s 的一个元，但 s 又包含全部普通集作为它的元，所以 s 不是普通集。这又得出一个矛盾，即 s 不是普通集。这是一个悖论；s 必然是普通集或非普通集，但每个可能性都导出了矛盾。

这个悖论并没有特别限制了集论。为了要弄清楚这点，我们将用一种没有多大意义的方法把这个悖论改造一下。这个作法完全脱离了集的概念。在一个兵团里，一个士兵被派作理发师工作，给他确切的命令是要他给团内的每个人修面，而团内规定每人都不准给自己修面。这个士兵是否应自己修面呢？如果他自己修了，而命令又不准他这样作。如果他没有自己修，则又违反了命令。这个人将如何严格执行命令呢？

这个悖论是纯逻辑性的。我们是不能避免把它引入集论中的，它显然比集论更一般，但集论的结构并不需要它。老的颇为呆板的逻辑主题已发展成为某些有趣的事情。事实上，数学家和逻辑学家已经同时工作了一些时候，要把逻辑从老的亚里士多德的形式中解放出来，但到现在很难看出将要采取怎样的新形式。

8. 一些组合问题

1. 一个简单例子可以说明我们将要研究的是怎样一种类型的问题。假设我们有四个红球(R)，一个黄球(Y)和两个白球(W)。假定这些球有同样的大小和重量，除其颜色

外，是完全不可分辨的。我们还假设有两个缸子A和B，缸子A恰好能装3个球，而B能装4个，我们问，把这七个有色球分别放在A和B两个缸内，将有多少不同的方法？

因为只有两个缸子，这是一个很简单的问题，我们只要看怎样把球放入A缸内，其余四个球就是放入B缸内的。为答复这一问题，我们将列举各种可能性。首先，我们把红球都放入A缸内，剩余的放入B缸内：

1. A缸内：RRR，B缸内：$RYWW$.

在这里哪三个红球放进A缸内是没有关系的，因为我们假设，除颜色外，球彼此间是没有差别的。

其次，A缸内放入两个红球，第三个球或是黄球或是白球。

2. A缸内：RRY，B缸内：$RRWW$，

3. A缸内：RRW，B缸内：$RRYW$.

如果A缸内只放入一个红球，则其他两球显然是YW或WW。

4. A缸内：RYW，B缸内：$RRRW$，

5. A缸内：RWW，B缸内：$RRRY$.

最后如果A缸内不放红球，则它将含有其他三个球：

6. A缸内：YWW，B缸内：$RRRR$.

所以，我们看到把4个红球，1个黄球和2个白球分别在A缸内放进3个，B缸内放进4个，有6种可能的方法。显然，球取什么颜色是无关紧要的；如果有4个黑球，1个绿球，2个蓝球，我们会得到同样的结果。需要说明的只是每种球的个数而不是它们的颜色。为了用一种稍为简单的形式说明这个问题，我们可以这样说：把各有4个，1个和2

个的 3 种球，分别装在两个容量为 3 个和 4 个的缸子内，共有 6 种分配方法。我们也可把这个事实用符号写成如下形式：

$$\{4,1,2|3,4\}_7=6,$$

其中直杠前的数字代表各种不同颜色的球数，而直杠后的数字代表每个缸内可容纳的球数。在直杠前的各数之和，必等于直杠后的各数之和。因所有的球必然全部分别放在缸子之内。球的全部个数 7 作为足标，已写在括号右边之下。

2. 问题并不是只需要考虑 3 种颜色和 2 个缸子而已。作为一般的问题，我们可以考虑有 c 种不同颜色的 n 个球，还有总容量为 n 个球的 u 个缸子。我们将使用符号

$$Z=\{r, s, \cdots|a, b\cdots\}_n, \qquad (1)$$

其中 n 个球分配的个数是，有 r 个是一种颜色，s 是另一种，等等，在缸子内分别含有 $a, b\cdots$ 个球。这个问题是要由数 $n, r, s, \cdots, a, b, \cdots$ 去计算 Z 的值。我们将不去解决这种一般形式的这个问题，而只限于研究一些例子和重要的特殊情况。

这个问题实际上并不要求分配的对象是色球。例如，我们曾用字母 $RRRRYWW$ 去代表 4 个红的，1 个黄的，2 个白的球。并把 7 个字母分配在两个集 A 和 B，以代替 7 个球分配在两个缸子 A 和 B。

3. 我们还可以选取一些并不涉及色球的例子，但我们将设法找出一种能解释每种情况的方法，就象对色球那样。这些例子是相当有趣而重要的，它们可用来说明符号 (1) 的意义。

例 I. n 个人在 n 个位置就座能有多少方式？为要把这个

解释成前一问题的形式，我们注意到，每个人将区别于其它任何人。因此，我们可以把每个人看作有不同的颜色，就如同每个人有不同的名字一样。那么这个问题就化为：n 个不同颜色的球，放在 n 个不同的位置有多少种放法？n 个位置中的每一个对应于老问题中的缸子，它将仅仅容纳一个球。用我们的符号，问题变为求

$$P_n=\{1,1,\cdots|1,1,\cdots\}_n \tag{2}$$

的值。这里，在直杠前的 n 个 1，对应于 n 个不同颜色的球，而在直杠后面的 n 个 1，对应于 n 个缸子，其中每个缸子都放进一个球。

如果我们设想这 n 个缸子或位置是在一行内，那么，我们要问 n 个球或人(或任何可分开的物体)在排成一行时有多少方法？这样的一个安排被叫作一个**排列**，所以我们可以说(2)中的 P_n 表示 n 个互异的物体的排列。

这里仍留有一个问题，如果已知 n 的值，那么怎样去求 P_n 的数值。我们将得到用 n 表示 P_n 的公式，但我们且等提出更多的例子之后，再解决这一问题。

4．例Ⅱ．3 个选手玩一种32张互异的纸牌游戏。每个人分到 10 张牌，剩下 2 张牌。按此要求把牌分完有多少种不同的方法？显然，这个分法数，就如有 32 个颜色不同的球放入容量分别为 10,10,10,2 的 4 个缸子内的分配方式的个数一样。所以，在这场纸牌游戏中，分配方法的个数是

$$S=\{1,1\cdots|10,10,10,2\}_{32}。 \tag{3}$$

仍然把对这个符号的数值计算推延一下。

5．例Ⅲ．所谓的多项式定理是又一有趣的例子。我们将仅仅考虑一个包含三个变数 x, y, z 的特例。下式

$$(x+y+z)^n$$

可以乘出来。这个 n 次方表示 n 个相同因子 $(x+y+z)$ 相乘的乘积。如果包含在括号内的各项之和被一因子所乘，则括号内每项都要被这个因子乘。如果这一规则都应用于这 n 次幂的 n 个括号，则可见这个 n 次幂由乘积的和组成。这些乘积每一个都有 n 个因子，它可以是 x 或 y 或 z。因为乘积的因子的次序可任意取，所以我们可以把所有的 x 集中在一起，同样对 y，对 z 也是如此。那么每一乘积将具有形式 $x^a y^b z^c$，其中 a, b, c 是整数，且受条件

$$a+b+c=n, \quad a\geqslant 0,\ b\geqslant 0,\ c\geqslant 0. \tag{4}$$

的限制。因子 x 可以来源于 n 个括号中的任一个，同样对于 y 和 z 也是如此。所以将有一些乘积 $x^a y^b z^c$ 有相同的 a, b, c。当我们乘出这个 n 次方时，将出现多少个 $x^a y^b z^c$ 呢？因子 x, y 或 z 之一，必定来自 n 个括号中的一个。我们可以设想 n 个缸子排成一行，每个缸子对应于一个括号。则我们可以放入每个缸子内一个字母，它是在对应的括号内取出的。现在我们只需要计算有多少种方法使 a 个元素"x"(红球)，b 个元素"y"(黄球)和 c 个元素"z"(白球)分配在容量为 1 的 n 个缸内。我们将以 $P_{a,b,c}^{(n)}$ 表示此数，我们有

$$P_{a,b,c}^{(n)}=\{a, b, c|1,1,\cdots,1\}_n。 \tag{5}$$

当 $(x+y+z)^n$ 乘出后，则项 $x^a y^b z^c$ 将出现 $P_{a,b,c}^{(n)}$ 次，这就是合并同类项后出现的系数。

一个特殊情况，将使得这一结果更清楚了。让我们取 $n=4$，且设想把四次幂 $(x+y+z)^4$ 乘出来。$x^a y^b z^c$ 型的项将会出现，其中 a, b, c 的所有的值应符合(4)，但其中 $n=4$。将它们系统地表列出来，我们求出下列 15 个可能性：

x^4, y^4, z^4,

x^3y, x^3z, xy^3, y^3z, xz^3, yz^3,

x^2y^2, x^2z^2, y^2z^2,

x^2yz, xy^2z, xyz^2。

把这许多项分别乘以各项的系数再相加，得

$$
\begin{aligned}
(x+y+z)^4 = & P^{(4)}_{4,0,0}\, x^4 + P^{(4)}_{0,4,0}\, y^4 + P^{(4)}_{0,0,4}\, z^4 \\
& + P^{(4)}_{3,1,0}\, x^3y + P^{(4)}_{3,0,1}\, x^3z + P^{(4)}_{1,3,0}\, xy^3 \\
& + P^{(4)}_{0,3,1}\, y^3z + P^{(4)}_{1,0,3}\, xz^3 + P^{(4)}_{0,1,3}\, yz^3 \\
& + P^{(4)}_{2,2,0}\, x^2y^2 + P^{(4)}_{2,0,2}\, x^2z^2 + P^{(4)}_{0,2,2}\, y^2z^2 \\
& + P^{(4)}_{2,1,1}\, x^2yz + P^{(4)}_{1,2,1}\, xy^2z + P^{(4)}_{1,1,2}\, xyz^2。
\end{aligned}
$$

这仅给出了乘方的形式。还必须求出$P^{(4)}_{a,b,c}$的值，更一般的是求$P^{(n)}_{a,b,c}$之值。

6．例Ⅳ．从n个不同物件中每次选取k个，有多少不同选法？这个数通常被说成是n个物件中每次选取k个的组合数，用来说明符号$C^{(n)}_k$。这n个物体是完全互异的，因而我们可以把它看作是不同颜色的球，即$r=1$，$s=1,\cdots$。所选取的k个球放在一个缸子内$a=k$，而其余都放在另一个缸子内，$b=n-k$，因而我们试图确定的数是

$$C^{(n)}_k=\{1,1,\cdots|k,n-k\}_n。\tag{7}$$

7．分配符号的对偶性。我们通常用以表示分配法的个数的符号有一重要特性，可以帮助我们去计算它的值。在直杠前和直杠后的数，可以互换，

$$\{r,s,\cdots|a,b,\cdots\}_n=\{a,b,\cdots|r,s,\cdots\}_n。\tag{8}$$

用进行分配的话说，就是把r个红球，s个白球…分别放进容量为$a,b,\cdots$个球的缸子内的分配方法的个数恰好等于把a个红球，b个白球…分别放进容量为$r,s,\cdots$个球的缸子

内的分法数。这里和通常一样，假定 $r+s+\cdots=a+b+\cdots=n$。

用这种分配方法来表示，等式(8)就很容易证明。一个简单的数值的例子，将完全证实这个结果。让我们证明等式

$$\{3,4|1,1,5\}_7=\{1,1,5|3,4\}_7。\qquad (9)$$

左边表示3个红球，4个白球的分配方法

$$R,R,R,W,W,W,W,$$

这里有3个缸子，A和B每个缸子放入1球，而C缸放入其它5球。我们看，可能的分配法之一是

$$\underset{A}{\boxed{R}}\qquad\underset{B}{\boxed{W}}\qquad\underset{C}{\boxed{RRWWW}}$$

现在我们也可以把这一分配法表示如下：把球写成一行，而在每个球之下，记下它将投入的缸子，

$$\begin{matrix}R,W,R,R,W,W,W,\\A,B,C,C,C,C,C,\end{matrix}\qquad(10\,a)$$

这正是把字母配成7对。上边的字母R和W是颜色的名称，下边的字母A,B,C是缸子的名称。我们可以互换字母所代表的东西，令A,B,C代表颜色的名称，而R,W则代表缸子的名称。现在我们把颜色写在上一行，缸子写在下一行，而依缸子的次序重新排列各对，得表

$$\begin{matrix}A,C,C,B,C,C,C,\\R,R,R,W,W,W,W,\end{matrix}\qquad(10b)$$

在(10 b)和(10 a)所出现的配对完全一样，只是恰好每对都颠倒过来了。但(10 b)可以理解为一种分配法，把1个A色的球，1个B色的球和5个C色的球放在两个缸子R和W

内，而 R 和 W 可分别放进 3 和 4 个球。但这正是(9) 式右边所算出的分配法之一。其他用(9)式左边所算出的每一个分配法，将对应于右边的一个，正如同样方法(10 a)对应于(10 b)一样。这一对应性显然可从另一方向作出，即以同一方法从(10 b)到(10 a)作出。所以这里有一个完全的对应性，即在(9)式左边和右边所表示的两个问题中，有一个"对偶"。由于分配法的集恰好配起对来，它们恰好有相同的数，因而(9)是一个真正的等式。

证明(9)式两边相等，不必知道两边的数值。但是在这一情况下，很容易把全部可能的分配法系统地列出表来，并得出

$$\{3,4|1,1,5\}_7=\{1,1,5|3,4\}_7=4。$$

这个(9)式的证法，仅是把颜色的名称和缸子的名称交换一下而已。我们不必去作(8)的详尽证明。颜色和缸子的同样的互换，表明了(8)式两端两个问题的对偶性，及其一般等式。

8. 某些情况下分配法数值的计算。对于任意数 $r,s,\cdots,a,b,\cdots$ 计算数值 $\{r,s,\cdots|a,b,\cdots\}_n$ 是十分麻烦的，我们不去作它。第 3 到第 6 节中所见的符号并不是最一般的类型。在直杠前的所有数字，或在其后的所有数字(或二者)都有特殊性，即全部由 1 组成。这就是说，或者所有的球都有不同的颜色，或者每一个缸子内只能放入一个球。由于对偶性，我们只须考虑全部的球都有不同颜色的情况，我们可以计算

$$\{1,1,\cdots,1|a,b,c,\cdots\}_n,$$

足标 n 提示我们这里有 n 个球，而全部缸子恰好共放入 n 个球，

$$1+1+\cdots+1=a+b+c+\cdots=n,$$

由于我们只考虑直杠前的符号 1，通常不需把它们写进去，我们可用简便记法

$$\{1,1,\cdots,1|a,b,c,\cdots\}_n=\{a,b,c,\cdots\}_n,$$

数$a,b,c,\cdots$表示缸子的容量，除其和为n外没有其他限制，显然，我们有

$$\{n\}_n=1, \tag{11}$$

因为把所有的球放在一个缸子内只有一种方法。如果把能容纳n个球的缸子N换成两个缸子，即容纳$n-1$个球的N_1和容纳1个球的N'。那么，必须从N中取出一个球放入N'内。余下的$n-1$个球放入N_1中。放入N'的那一个球可以是任意一个，因为它们是互异的，从而有n个可能性。所以，我们有

$$\{n-1,1\}_n=n,$$

同样方法我们证明

$$a\{a,b,c,\cdots\}_n=\{a-1,1,b,c,\cdots\}_n。 \tag{12}$$

这里具有容量$b,c,\cdots$的缸子$B,C,\cdots$没有改变，但容量a的缸子A为两个缸子A_1和A'所代替，它们的容量分别为$a-1$和1个球。我们可以从$A,B,C,\cdots$中的一个分配法转到对$A_1,A',B,C,\cdots$的一个分配法，就是从A中取任一球放入A'内，而将其余$a-1$个球放入A_1中，而$B,C,\cdots$中的球不变。还有，因为球都是互异的，放入A'中的球有a个不同的可能选法，所以对于$A_1,A',B,C\cdots$的分配个数，应是对于$A,B,C,\cdots$的分配个数的a倍，而这恰是(12)的意义。

现在我们可以把A_1换成容纳$a-2$个球的缸子A_2和容

纳1个球的缸子 A''，且显然有

$$(a-1)\{a-1,1,b,c,\cdots\}_n=\{a-2,1,1,b,c,\cdots\}_n。$$

同一方法得到

$$(a-2)\{a-2,1,1,b,c,\cdots\}_n=\{a-3,1,1,1,b,c,\cdots\}_n,$$

如此，得到一组全部类似的方程，其最后一个是

$$2\{2,\underbrace{1,1,\cdots,1}_{a-2},b,c,\cdots\}_n=\{\underbrace{1,1,\cdots,1}_{a},b,c,\cdots\}_n$$

最后这一步，已把 a 分裂成一组1了。如果我们把(12)和它后面的所有方程乘到一起，两端都有公因子

$$\{a-1,1,b,c,\cdots\}_n,\{a-2,1,1,b,c\cdots\}_n,\cdots,$$
$$\{2,\underbrace{1,1,\cdots 1}_{a-2},b,c,\cdots\}_n,$$

它们可以从两端消去。用这一方法，我们求得

$$a(a-1)(a-2)\cdots 2\{a,b,c,\cdots\}_n=\{\underbrace{1,1,\cdots,1}_{a},b,c,\cdots\}_n。$$

在乘积 $a(a-1)(a-2)\cdots 2$ 中颠倒其次序，且引用符号

$$a!=1\cdot 2\cdot 3\cdots(a-1)\cdot a,$$

我们的结果可以写成

$$a!\{a,b,c,\cdots\}_n=\{\underbrace{1,1,\cdots 1}_{a},b,c,\cdots\}_n。\qquad(13)$$

可以采用完全同样的步骤把缸子 B 换成较小的一些缸子。这给出

$$b!\{\underbrace{1,1,\cdots,1}_{a},b,c,\cdots\}_n=\{\underbrace{1,1,\cdots,1}_{a},\underbrace{1,1,\cdots 1}_{b},c,\cdots\}_n$$

这和(13)结合起来，得

$$a!b!\{a,b,c,\cdots\}_n=\{\underbrace{1,1,\cdots,1}_{a+b},c,\cdots\}_n,$$

同法处理 c，并连续处理其余的缸子，我们最后得到

$$a!\,b!\,c!\cdots\{a,b,c,\cdots\}_n=\{1,1,\cdots,1\}_n, \tag{14}$$

其中，显然右边的符号中恰好包含 n 个 1。

在(13)内，如果 $a=n$，这时只有一个缸子，而(13)变为

$$n!\{n\}_n=\{1,1,\cdots,1\}_n,$$

所以，用(11)得到

$$n!=\{1,1,\cdots,1\}_n。 \tag{15}$$

用这个值代入(14)，我们得到公式

$$\{a,b,c,\cdots\}_n=\{1,1,\cdots,1|a,b,c,\cdots\}_n=\frac{n!}{a!\,b!\,c!\cdots}。 \tag{16}$$

利用这个符号的计算式，至少上述全部的例子都可以得到解决。

9. 现在我们可以回到特殊的例子

例 I　从(2)和(15)，我们有

$$P_n=\{1,1,\cdots,1|1,1,\cdots,1\}_n=\{1,1,\cdots,1\}_n=n!$$

就是说，n 个人分别坐在 n 个位子上，有 $n!$ 个不同坐法，或者说 n 个不同物件有 $n!$ 个排列，在这一公式中所出现的阶乘数值增加得很快。前边十个是

1!=	1	6!=	720
2!=	2	7!=	5,040
3!=	6	8!=	40,320
4!=	24	9!=	362,880

$$5!=120 \qquad 10!=3,628,800$$

例Ⅱ　从(3)和(6)可见

$$S=\frac{32!}{10!10!10!2!}。$$

这就是在这一种游戏中，32张牌的不同分配方法的个数。实际上这是一个很大的数，计算的结果是

$$S=2,753,294,408,504,640。$$

在打桥牌时，读者有兴趣的话可用类似的方法加以计算。

例Ⅲ　利用对偶性公式(8)以及(5)和(16)，我们有

$$P_{a,b,c}^{(n)}=\frac{n!}{a!\,b!\,c!} \quad (a+b+c=n).$$

如对 $n=4$ 加以计算，得展开式

$$\begin{aligned}(x+y+z)^4 &= x^4+y^4+z^4\\ &+4x^3y+4x^3z+4xy^3+4y^3z+4xz^3+4yz^3\\ &+6x^2y^2+6x^2z^2+6y^2z^2\\ &+12x^2yz+12xy^2z+12xyz^2。\end{aligned}$$

例Ⅳ　最后(7)和(16)得出值

$$C_k^{(n)}=\frac{n!}{k!(n-k)!}。$$

这是 n 个物体中每次取 k 个的组合数。

9. 关于韦林问题

平方数的序列 1,4,9,16,25，…，当这样不断列下去时，两个相邻的平方数的间隔越来越大，这就是说，越来越

不稠密了。虽然这些间隔中的数都不是平方数，但至少有些数可以写作两个平方数之和，例如 13＝9＋4，41＝25＋16，等等。但不是每个数都可写作两个平方数之和。如果我们试图把 6 写成两个平方数之和，可用的平方数是 1 和 4，这是小于 6 的仅有的两个平方数。我们看 1＋1，4＋4 或 1＋4 都不能等于 6，所以为使和得 6，至少需要三个平方数。事实上，6 可以写作三个平方数之和，6＝4＋1＋1。同样的方式，可以看出 7 不能写作三个平方数之和，因为最少的个数是四，即 7＝4＋1＋1＋1，对于 8，两个平方数就够了，8＝4＋4，9 自己就是一个平方数。又我们有 10＝9＋1，11＝9＋1＋1，12＝9＋1＋1＋1＝4＋4＋4，等等。

我们自然会认为不久四个平方数将是不够的了，因为这样继续下去，将需要更多的平方数。可是，费马，这个在17世纪同笛卡尔并列的大数学家，证明了令人惊异的事实：每个正整数最多可以表示为四个平方数之和。

韦林推测，类似对立方数，四次方数等等而提出的问题，即应需要多少个立方数，四次方数等等的问题，也是可以证明的。由此，他的名字就和这一系列问题联系在一起了。立方数是 1，8，27，64，…，如果我们试图把不大的数表示成较它小的立方数之和，例如 7，这个 8 前面的最后一个数，就必须全部用一些 1 表示出来，7＝1＋1＋1＋1＋1＋1＋1，需要七个立方数。同样，15＝8＋1＋1＋1＋1＋1＋1＋1，需要八个立方数。而 23＝8＋8＋1＋1＋1＋1＋1＋1＋1，需要九个。在快达到 31 之前，我们发现一个新立方数 27 可以利用，从而全部情况改变了。事实上，31＝27＋1＋1＋1＋1，仅用五个立方数。

雅各比有一位帮他计算的名叫达斯的人，编出一个表来，表中每个数都表示成若干个最少的立方数之和。这个表表明，在 23 之后，下一个用九个立方数表出的数是 239。这是在这个表中一直到 12000 时仅有的一个用九个立方数表出的数。恰需用八个立方数表达的数有 15,22,50,114,167,175, 186, 212,213,238,303,364,420,428, 454, 此后， 一直到 12000 都没有这样的数出现。需要七个立方数的有 7,14,21,42, 47, 49,61,77、85,87,103,…, 5306, 5818, 8042, 但对这一序列，表明已到了尽头。继续这种单凭经验的工作，我们只是反复地证实了这些事实。

这类经验工作不能证明任何东西。它只能有助于指出这可能是真实的，即每个数最多是九个立方数之和，或从某个定数起，每个数最多是八个或者是七个立方数之和。后一命题，即由某数起用八个立方数就够了，这是首先由蓝多用很难的数字方法证明的。此后逐步发展，威弗里茨证明了前一命题。

四次方数可用对立方数的同样方式去处理它。首先，一些四次方数是 1, 16, 81, 256, …, 现在 15 需要 15 个四次方数，31 要 16 个。47 要 17 个，63 要 18 个，79 要 19 个。此后加入了新四次方数，而形式完全变了。现在的问题是，19 个四次方数是否足够用了？对这一问题曾做过很多工作。柳维尔证明 53 个足够了。 这个数逐渐被压缩为 47,45,41,39, 38, 而后威弗里茨得到 37。 不过这和所希望的 19 还差很远。

德国大数学家希尔伯特着手用另外的方法解决这个一般问题。他不是试图改进以前的结果，而是总的考虑与立方数、

四次方数等等有关的全部一系列问题。他能一举证明，不但对立方数和四次方数，而且对五次方，六次方以及多次方数，总有一个数是足够的（象 9 和 37 对于立方数和四次方数是足够的）。显然，对于高次方数这个数要较大些。

英国的哈迪和利特尔伍德使用另外的更复杂的方法去解决这个问题。他们证明了从某一数起全部的数可以写作 19 个四次方数。哈迪和利特尔伍德的结果说明，有这样一个数N，从N开始全部的数一定最多是 19 个四次方数之和。但这个数N，由他们的证明中得知是非常大的，从而使他们不必实际地计算出它的值来。以四次方数而言，需要逐个地检验小于N的每个数，看它们是否用 19 个四次方数就足够了。不过，数N是很大的，这样的计算将大大超过任何计算者的能力了。

我们曾用了不少的篇幅讨论了联系这一问题的事实。这种讨论将给出一个概念，即经验研究是怎样用以帮助发现事实而发展新的理论的。现在我们试图给出联系这个问题的方法的一些概念，特别是希尔伯特所采用的。很遗憾，哈迪和利特尔伍德的富有权威的方法是太先进而复杂了，以致未能纳入，甚至我们将提到的一些证明，也部分地超出了本题的讨论范围，但说出所用方法的概念还是可能的。

象通常那样，我们从简单的情况开始。方程$(a+b)(a-b)=a^2-b^2$是代数中常见的。假如忘记了，把左边乘起来就很容易验明。这个方程无论a和b是什么数都是正确的。一个永远正确的方程式，我们把它叫作“恒等式”。更复杂一点的恒等式是

$$(a^2+b^2)(c^2+d^2)=(ac+bd)^2+(ad-bc)^2。\quad (1)$$

为要验明此式，我们记起公式

$$(x+y)^2=x^2+2xy+y^2,$$

用它把右边乘出来，得

$$(a^2c^2+2abcd+b^2d^2)+(a^2d^2-2abcd+b^2c^2)$$
$$=a^2c^2+a^2d^2+b^2c^2+b^2d^2+2abcd-2abcd。$$

末尾两项彼此消去了，其它可合并，得出 $a^2(c^2+d^2)+b^2(c^2+d^2)=(a^2+b^2)(c^2+d^2)$这正是(1)式左边所列的。

这个恒等式可产生一些有趣的结果：如果两数中每一数都是两个平方数之和，则其乘积也是两个平方数之和。例如 $13=9+4$，又 $41=25+16$，二者都是这种形式，则，依(1)我们有

$533=13\cdot41=(3^2+2^2)\ (5^2+4^2)=(3\cdot5+2\cdot4)^2+(3\cdot4-2\cdot5)^2=23^2+2^2$，我们得到 533 可以表作两个平方数之和。公式(1) 可按同法用于任何两数，而它们都是两个平方数之和的情况。

欧拉是 18 世纪瑞士的大数学家，发现了下列恒等式

$$(a_1^2+a_2^2+a_3^2+a_4^2)(b_1^2+b_2^2+b_3^2+b_4^2) \qquad (2)$$
$$=(-a_1b_1+a_2b_2+a_3b_3+a_4b_4)^2+(a_1b_2+a_2b_1+a_3b_4-a_4b_3)^2$$
$$+(a_1b_3-a_2b_4+a_3b_1+a_4b_2)^2+(a_1b_4+a_2b_3-a_3b_2+a_4b_1)^2.$$

如果两边都乘出来，再应用熟知的公式，

$$(x_1+x_2+x_3+x_4)^2=x_1^2+x_2^2+x_3^2+x_4^2+2x_1x_2$$
$$+2x_1x_3+2x_2x_3+2x_1x_4+2x_2x_4+2x_3x_4。$$

这个恒等式可以毫无困难地加以验明。

公式(2)类似于(1)，它表明：如有两个数，其中每个都是四个平方数之和，则其乘积也是四个平方数之和。拉格朗日应用这个恒等式对费马定理作出了一个很出色的证明。这

个定理说：每个正整数可以表示成最多是四个平方数之和。首先，可作这样一个注释，即仅仅需要证明每个素数可以表示成四个平方数之和就可以了，因为这就自然地推到合成数了。讨论素数时，拉格朗日还是应用(2)。目前，我们且先接受费马定理，把它当作已经成立，先使用它。拉格朗日对它的证明将在本篇最末给出。

柳维尔利用这个定理来证明每个数最多是53个四次方数之和。他还用了恒等式：

$$
\begin{aligned}
&6(x_1^2+x_2^2+x_3^2+x_4^2)^2\\
&=(x_1+x_2)^4+(x_1+x_3)^4+(x_2+x_3)^4+(x_1+x_4)^4+\\
&(x_2+x_4)^4+(x_3+x_4)^4+(x_1-x_2)^4+(x_1-x_3)^4+ \qquad (3)\\
&(x_2-x_3)^4+(x_1-x_4)^4+(x_2-x_4)^4+(x_3-x_4)^4。
\end{aligned}
$$

要验明上式，首先应用二项式定理，展开

$$(x_1+x_2)^4=x_1^4+4x_1^3x_2+6x_1^2x_2^2+4x_1x_2^3+x_2^4,$$

$$(x_1-x_2)^4=x_1^4-4x_1^3x_2+6x_1^2x_2^2-4x_1x_2^3+x_2^4,$$

把这些加起来，得 $(x_1+x_2)^4+(x_1-x_2)^4=2x_1^4+2x_2^4+12x_1^2x_2^2$,

对于(3)式中第二行每个括弧和它下面的那两个括弧相加求出(3)的右边，有展开式

$$
\begin{aligned}
&6(x_1^4+x_2^4+x_3^4+x_4^4)+12(x_1^2x_2^2+x_1^2x_3^2\\
&\quad+x_2^2x_3^2+x_1^2x_4^2+x_2^2x_4^2+x_3^2x_4^2)。
\end{aligned}
$$

现在如果把(3)式左边也乘出来，立刻可见有相同的值。

柳维尔在下述方法中使用了这个恒等式。令 n 为任一正整数，证明 n 必然最多是53个四次方数之和。他先把 n 除以6，再求商数和余数，$n=6x+y$(如果 n 是29，则商 x 是4，余数是5)，这里 y 将是0,1,2,3,4,5中之一。在这一点上，

柳维尔第一次应用了费马定理，他应用这个定理把 x 写作四个平方数之和，$x=a^2+b^2+c^2+d^2$。这时，原数可以写作

$$n=6x+y=6(a^2+b^2+c^2+d^2)+y$$
$$=6a^2+6b^2+6c^2+6d^2+y。$$

对于 a,b,c,d，柳维尔现在再次使用费马定理，得

$$a=a_1^2+a_2^2+a_3^2+a_4^2,$$
$$b=b_1^2+b_2^2+b_3^2+b_4^2,$$
$$c=c_1^2+c_2^2+c_3^2+c_4^2,$$
$$d=d_1^2+d_2^2+d_3^2+d_4^2,$$

所以 $$n=6(a_1^2+a_2^2+a_3^2+a_4^2)^2+\cdots+$$
$$6(d_1^2+d_2^2+d_3^2+d_4^2)^2+y。$$

现在可以应用恒等式(3)。应用到右边第一项，断定此项可以写作 12 个四次方数之和。同样对于其它类似项也是这样。因此，$4\cdot12=48$ 四次方数已经被应用，而 y 仍可分成四次方数。因为 y 是 0,1,2,3,4 或 5，它可以表示为最多是五个四次方数之和，其中每个四次方数都是 1。这样给出全部 $48+5=53$ 个四次方数。

拉格朗日定理是说，每个正整数N可以写作四个平方数之和，$N=z_1^2+z_2^2+z_3^2+z_4^2$。为证明这一定理，我们将应用本篇的公式(2)，但要改变其字母。如果我们令 $a_1=x_1$, $a_2=x_2$, $a_3=x_3$, $a_4=x_4$, $b_1=-y_1$, $b_2=y_2$, $b_3=y_3$, $b_4=y_4$ 我们得出

$$\begin{aligned}&(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2)\\&=(x_1y_1+x_2y_2+x_3y_3+x_4y_4)^2+(x_1y_2-x_2y_1\\&+x_3y_4-x_4y_3)^2+(x_1y_3-x_3y_1+x_4y_2-x_2y_4)^2\\&+(x_1y_4-x_4y_1+x_2y_3-x_3y_2)^2。\end{aligned}\tag{2'}$$

证明可以分成若干步，我们先证明

定理 1. 如果 A 和 B 都可以写成四个平方数之和，则乘积 AB 也是这样。这可从(2) 直接得到，因为如果 $A=x_1^2+x_2^2+x_3^2+x_4^2$ 又 $B=y_1^2+y_2^2+y_3^2+y_4^2$，则(2)表出 AB 是四个整数的平方之和。这个结果将使我们集中注意于素数，因为它说明我们仅需考虑一个数的素数因子。我们不打算立即证明对于素数的全部结果，但将开始考虑：

定理 2. 如果 p 是大于 2 的一个素数，则可能求出一个整数 m 使得 $1\leqslant m<p$，而使 mp 可以写成四个平方数之和，$mp=x_1^2+x_2^2+x_3^2+x_4^2$。 如果不坚持 $1\leqslant m<p$，这是很容易证明的。因为我们可以写成 $0\cdot p=0^2+0^2+0^2+0^2$ 或 $p\cdot p=p^2+0^2+0^2+0^2$ 所以需要 $1\leqslant m<p$，这在以后将是重要的。

试证明定理 2，首先注意 p 是一奇数，因它是素数而又大于 2。我们写出 $0^2, 1^2, 2^2, \cdots, \left(\frac{p-1}{2}\right)^2$，每个除以 p，而仅保留余数。这样，给出 $\frac{p+1}{2}$ 个余数 r，每个在 0 与 $p-1$ 之间。设 $p=11$，我们将写出 0,1,4,9,16,25。每个除以 11，得出余数 r 是 0,1,4,9,5,3。在各种情况下，全部余数将是互异的。如果其中两个相等，我们将有两个在 0 和 $\frac{p-1}{2}$ 之间的整数，$x_1>x_2$，其平方除以 p 时将具有同一余数 r。就是说 $x_1^2=q_1p+r$ 又 $x_2^2=q_2p+r$，相减得 $x_1^2-x_2^2=(q_1-q_2)p$ 或 $(x_1-x_2)(x_1+x_2)=(q_1-q_2)p$。因为 p 是素数，它必然要整除 x_1-x_2 或 x_1+x_2，但这是不可能的，因 x_1-x_2 和 x_1+x_2 是小于 p 的正数。

现在我们取余数 r，每个都加 1，再由 p 减去这些数。这样给出 $\frac{p+1}{2}$ 个在 0 与 $p-1$ 之间的数 s，且它们是互异的。对于 $p=11$，我们将得到 10,9,6,1,5,7。s 中至少有一个数等于前组 r 中的一个数，因为 r 在 p 个数 0，1，2，… $p-1$ 中仅使用了 $\frac{p+1}{2}$ 个，剩下的仅有 $\frac{p-1}{2}$ 个数，而同时有 $\frac{p+1}{2}$ 个 s。对于 $p=11$，我们看到数 1,9,5 都在 r 和 s 之中。

令 $R=S$ 为 r 和 s 中相等之数。R 是某个 x^2 除以 p 所剩的余数，其中 $0\leqslant x\leqslant\frac{p-1}{2}$。$S$ 是某一数 y^2 除以 p 后的余数加 1，再从 p 中减去，其中 $0\leqslant y\leqslant\frac{p-1}{2}$。这就是 $x^2=q_1p+R$，$y^2=q_2p+r$，$S=p-(r+1)$。把以上三式相加，得 $x^2+y^2+S=(q_1+q_2+1)p+R-1$。因为 $R=S$，我们可以写成 $x^2+y^2+1=mp$，其中 $m=q_1+q_2+1$。还有因 $0\leqslant x\leqslant\frac{p-1}{2}$ 又 $0\leqslant y\leqslant\frac{p-1}{2}$，我们有 $0<mp\leqslant\left(\frac{p-1}{2}\right)^2+\left(\frac{p-1}{2}\right)^2+1=\frac{p^2-2p+1}{2}+1=\frac{p^2-2p+3}{2}<\frac{p^2}{2}<p^2$，因而 $0<m<p$。这就证明了定理 2，因为我们有 $1\leqslant m<p$ 和 $mp=x^2+y^2+1^2+0^2$。

如果 $p=11$，我们可取 $R=S=5$，求出 $x=4$，$y=4$，$4^2+4^2+1^2+0^2=3\cdot 11$。不过如果取 $R=S=1$，则求得 $x=1$，$y=3$，$1^2+3^2+1^2+0^2=1\cdot 11$。在这种情形下，可把 $p=11$ 写作四个平方数之和。我们现在所用的方法，不是总能给出

p 是四个平方数之和的，不过我们可以证明：

定理 3. 如果 p 是大于 2 的一个素数，又假设 m 是使 mp 是四个平方数之和的最小的正整数，则 $m=1$。从定理 2 我们已知 $m<p$。这个最小的 m 值不能是偶数，因为假设是偶数，则 $x_1^2+x_2^2+x_3^2+x_4^2=mp$，将是一个偶数，所有的 x 或都是偶数，或二偶二奇或都是奇数。对于第二种情形，可假设 x_1, x_2 是偶，而 x_3, x_4 是奇。对每一种情形 (x_1+x_2)，(x_1-x_2)，(x_3+x_4)，(x_3-x_4) 就都是偶数

$$\left(\frac{x_1+x_2}{2}\right)^2+\left(\frac{x_1-x_2}{2}\right)^2+\left(\frac{x_3+x_4}{2}\right)^2+\left(\frac{x_3-x_4}{2}\right)^2$$

$$=\frac{1}{2}(x_1^2+x_2^2+x_3^2+x_4^2)=\frac{m}{2}p,$$

在左边全部数都是平方数，因而是整数。就是说 $\frac{m}{2}p$ 可以写作四个平方数之和，而偶数 m 将不是最小可能的数，而这是我们所假定的。

现在知道最小的 m 是奇数，又 $m<p$。要证明 $m=1$。我们假设 $m>1$，然后证明，它可再简化。因 m 是奇数，我们可假设 $m\geqslant 3$，则有

$$mp=x_1^2+x_2^2+x_3^2+x_4^2, \tag{4}$$

用 m 除每个 x_K 得余数 r_K，$0\leqslant r_K<m$。如果 $0\leqslant r_K\leqslant\frac{m-1}{2}$，我们写作 $y_K=r_K$。如果 $\frac{m+1}{2}\leqslant r_K\leqslant m-1$，则写作 $y_K=r_K-m$。在两种情形下，我们有 $x_K=q_K m+y_K$，以及 $-\frac{m-1}{2}\leqslant y_K\leqslant\frac{m-1}{2}$。由于 $y_K=x_K-q_K m$，我们还有，引用(4)

$$y_1^2+y_2^2+y_3^2+y_4^2=x_1^2+x_2^2+x_3^2+x_4^2$$
$$-2m(x_1q_1+x_2q_2+x_3q_3-x_4q_4)$$
$$(5)\qquad +m^2(q_1^2+q_2^2+q_3^2+q_4^2)=mp-2m(x_1q_1+x_2q_2$$
$$+x_3q_3+x_4q_4)+m^2(q_1^2+q_2^2+q_3^2+q_4^2)=mn$$

其中 n 是一整数。更进一步，我们有 $n>0$，因为如果$n=0$，我们将有 $y_1=y_2=y_3=y_4=0$。这意味着每个 x 都是 m 的倍数，即每个 x^2 都是 m^2 的倍数。从(4)我们看 mp 将是 m^2 的倍数，因而 p 是 m 的倍数。但这是不可能的，因为 p 是素数，而且 $1<m<p$。我们还有 $mn=y_1^2+y_2^2+y_3^2+y_4^2\leqslant 4\left(\frac{m-1}{2}\right)^2<m^2$因而 $n<m$。

把(4)和(5)乘起来，求出

(6) $m^2np=(x_1^2+x_2^2+x_3^2+x_4^2)(y_1^2+y_2^2+y_3^2+y_4^2)=(2')$ 式的右边。$(2')$式右边所出现平方数的第一个数是

$$\begin{aligned}
&x_1y_1+x_2y_2+x_3y_3+x_4y_4\\
&\quad=x_1(x_1-q_1m)+x_2(x_2-q_2m)+x_3(x_3-q_3m)\\
&\quad\quad+x_4(x_4-q_4m)\\
&\quad=x_1^2+x_2^2+x_3^2+x_4^2-m(x_1q_1+x_2q_2+x_3q_3+x_4q_4)\\
&\quad=mp-m(x_1q_1+x_2q_2+x_3q_3+x_4q_4)\\
&\quad=mz_1,
\end{aligned}$$

此处 z_1 是一整数。$(2')$式右边第二个数是

$$\begin{aligned}
&x_1y_2-x_2y_1+x_3y_4-x_4y_3\\
&\quad=x_1(x_2-q_2m)-x_2(x_1-q_1m)+x_3(x_4-q_4m)\\
&\quad\quad-x_4(x_3-q_3m)\\
&\quad=m(-x_1q_2+x_2q_1-x_3q_4+x_4q_3)\\
&\quad=mz_2,
\end{aligned}$$

这里 z_2 也是一个整数。同样方法，可以求出第三项和第四项 mz_3 和 mz_4。把这些放入(6)内，我们有

$$m^2np=m^2z_1^2+m^2z_2^2+m^2z_3^2+m^2z_4^2$$

$$np=z_1^2+z_2^2+z_3^2+z_4^2。$$

因此 np 可以写作四个平方数之和，而且我们早已求得 $0<n<m$。这证明了 $m>1$ 不是最小可能数值，如象我们曾经假设的那样。所剩下的只有 $m=1$，而定理 3 得到证明。

对于 $p=11$ 时，$x_1=4$，$x_2=4$，$x_3=1$，$x_4=0$，$m=3$，我们有 $y_1=1$，$y_2=1$，$y_3=1$，$y_4=0$；$4\cdot1+1\cdot4+1\cdot1+0\cdot0=3z_1$，$z_1=3$，$4\cdot1-4\cdot1+1\cdot0-0\cdot1=3z_2$，$z_2=0$；$4\cdot1-1\cdot1+0\cdot1-1\cdot0=3z_3$，$z_3=1$；$4\cdot0-0\cdot1+4\cdot1-1\cdot1=3z_4$ $z_4=1$，最后 $3^2+0^2+1^2+1^2=1\cdot11, n=1$。

定理 4. 如果 p 是一素数，它可以写作四个平方数之和。这几乎只是把定理 3 重述一遍。如果 $p=2$，我们有 $2=1^2+1^2+0^2+0^2$，定理是对的。如 $p>2$，则由定理 3，此定理也是对的。

最后，我们要谈拉格朗日的定理：

定理 5. 每个正整数 $n\geqslant0$ 可以写作四个平方数之和。这对于 $n=0$ 和 $n=1$ 是正确的，因为我们有 $0=0^2+0^2+0^2+0^2$，又 $1=1^2+0^2+0^2+0^2$。如果 n 是一素数，则由定理 3，知此定理也是正确的。剩下的只是对于合成数。如果 n 是一合成数，我们可以分解成素因子的乘积，$n=p_1p_2p_3\cdots p_t$，其中 $p_1, p_2, p_3, \cdots p_t$ 是素数，且不必是互异的。现在由定理 4，p_1 和 p_2 可以写作四个平方数之和，则由定理 1，p_1p_2 也可写作四个平方数之和。再由定理 4，p_3 可以写作四个平方数之和，再用定理 1，其乘积 $p_1p_2p_3$ 也可以写作四个平方数

之和。继续这样下去，最终我们发现，任一正整数 n 都可以写作四个平方数之和。

10. 关于闭自交曲线

1. 本篇讨论的是一种特殊类型的曲线。虽然这种曲线可以很复杂，然而都必须满足一定的条件。第一，它们按照确定的路线是走得通的。即从某点开始，铅笔一直不离开纸，可以一笔把整个曲线画完。第二，它们必须是封闭的。即当某人画曲线时，他应能够从已知点开始，走遍整个曲线，然后回到原来的起始点，并正好把曲线画出来。最后，曲线本身可以相交任意多次，但曲线通过每一个交点只有二次。例如图 26 和图 27 所示曲线就能满足这些条件，而图 28 就不能。图 26 和图 27 中所表示的那种类型的交点是允许的，叫



图 26

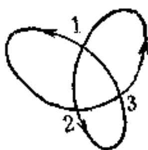


图 27


图 28

作二重点。图 28 所示的那种类型的交点是不许可的，叫作三重点。[1]当整个曲线走完一遍时，显然每个二重点都要经过二次。如果我们用数字来标明每个二重点时，那么就可以用一系列数字来表示所经过的二重点的次序。例如，图 26

有 1221 这样的次序，而图 27 有 123123。因为每个二重点都要经过二次，每一个数字在数列中必出现二次。高斯指出，认为每个数字都出现二次的任何一个数列都可以表示某曲线上二重点的想法是不正确的。在有二个二重点的情形时，我们已经遇到了 1221 这样的次序，但是，次序为 1212 的曲线是没有的。对于这种十分简单的情形，可以通过试验和对比加以验证。

这一节的主要结果将是这样一个定理，即在（表示二重点的次序的）一个数列中，每个二重点在偶数位出现一次，在奇数位也出现一次。这个说法可以略加改变，即我们断定在每个二重点出现的二次之间，或是相隔偶数位，或是相互紧挨着的。由这个定理可以马上看出，1212 是不可能的，因为二个 1 之间相隔只是一位。

2．为了证明这个定理，我们考虑曲线 A（图 29）上的任意一个二重点 Q。如果我们由 Q 开始，并且顺着曲线 A 走，我们将最终回到 Q。当第一次回到 Q 时，我们已经走过了整个曲线 A 的一部分 B。因为曲线由 Q 发出放射状的四条线段，而我们走过的只是二段，其中有离开 Q 的一段和回到 Q 的一段，所以这部分只能是曲线的一部分。如果我们继续通过 Q 沿曲线 A 走下去，我们就将走完 A 的剩余的部分 C。B 和 C 都是闭曲线，每条曲线在 Q 都有一个尖角。虽然 Q 是 A 的一个二重点，但不是 B 或 C 的二重点（每条曲线在此都不相交），并且 B 和 C 在此处连接而不相交。我们必须证明，沿 B 由 Q 回到 Q 时，经过二重点的次数必为偶数[2]。注意，在 B 上的 A 的二重点，或是 B 自身相交的点（B 的二重点），或是 B 和 C 相交的点。

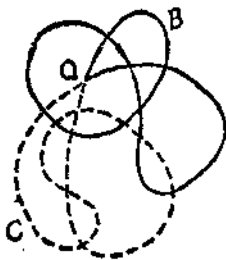


图 29

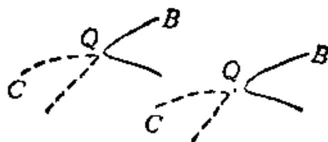


图 30

当沿B行走时，通过B的每个二重点必有二次。因而所有这些二重点合在一起，我们得出了偶数的记数次数。在B和C的交点处，是二个线段相交，其中一段是B的一部分，另一段是C的一部分。在沿B行走时，没有遇到属于C的那部分，那么经过B和C的交点只有一次。现在我们只需要证明B和C的交点只能是偶数个。

3．我们可以采取使B和C断开的办法(图 30)，同时要使邻近Q点的曲线稍微变形而不改变B和C在其它点的相交性。现在二条曲线可以在许多点彼此相交，但除了交点外，二条曲线再没有互相连接的地方了。

现在，我们要证明二条这样的曲线或者不相交，或者有偶数个点相交。为了证明这点，我们将把两个曲线中的一个，假设为C，逐步变形。在每一步，我们都将使C去掉一个二重点，因而使之适当的简化。如同在Q点一样，我们总可以在不破坏B和C的相交性的情况下，把曲线稍微变形。

设P是C的二重点。如同A在Q点分解为B和C那样，

曲线C在P点也可分解为二条闭曲线D和E。曲线D和E在P点彼此连接。如果我们在C上标出一定的方向作为沿C行走的方向，那么D和E就自动地给出了方向(图 31 a)。现在，

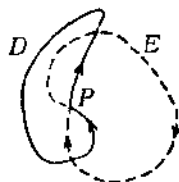


图 31a


图 31b



图 31c

由P开始，我们按标出的方向沿D行走，最后回到P。然后与所标的方向相反，沿E行走回到P。由于利用E的反方向，使我们去掉了在P处的交点。我们有了一个确定的路径来沿C行走，即需要经过P二次，并有二条在P为尖角的路径，但这二条路径在此不相交。在P彼此连接的二条路径，我们可以把它们稍微分开，并且把尖角弄圆。然后我们用在P处没有二重点的曲线代替C(图 31 b)。现在的曲线就比原来的曲线C少了一个二重点。当使曲线变形接近P时，我们必须做得很仔细，使C的其它二重点以及B和C的相交性不受破坏。

对于每个二重点，我们可以重复上述全部程序，最后就变成没有二重点的曲线C^*(图 31 c)。二条曲线C和C^*是几乎相同的。它们的差别只表现在C的二重点附近。对于我们来说，重要的是这二条曲线与B有相同的交点。

4. 一条没有二重点的闭曲线所包围的区域，我们叫作曲线的“内部”。这个事实看起来是颇为明显的，我们在直观

的基础上将承认它。平面上的既不在直线上，又不是曲线内部的那部分，叫作“外部”，并且由曲线本身把它与内部分隔开。如果这样的曲线与另外的曲线在某些点相交，那么这第二条曲线必然经过这些点，并且是由内部到外部，或相反是由外部到内部的。

我们现在可以证明，曲线B和C或者没有交点，或者相交点是偶数个。我们可以用没有二重点的C^*来代替C而不改变相交性。可能碰巧B完全在C^*的内部，在此情况下它们不相交。如果B完全在C^*的外部，它们仍然不相交。最后，B可能一部分在C^*的外部，一部分在C^*的内部。在这种情况下，令T是在C^*外的B上的任一点。如果我们由T开始并沿B行走，必然会首次进入C^*内部。这就是B和C^*的一个交点。因为B是封闭的，我们必然会最终回到C^*外部的T；即我们必然由内部到外部使又一次相交。可能B进入内部有好几次，但B每进入内部一次，必随着B要离开内部一次，因为B最后是要回到外部的起点T的。因此，在所有情况下，C^*和B（并且因此C和B）或者相交为偶数次，或者完全不相交。正如第二节末尾所指出的，这对于定理的完整证明已经足够了。

5．这个定理的关于A的每个二重点在偶数位出现一次，且在奇数位出现一次的结论，可以用另一种不同的方式解释。我们设想A是空间作出的一条曲线的投影，用A的二重点表示空间曲线的一部分经过另一部分的地方的投影。如果曲线是一条道路，那么二重点就表示立体交叉路口。我们希望这样安排路径，即当沿它行走通过每个交叉口时，是交错的：一次由上面过，另一次由下面过。问题是，这种情况总

是存在的吗？如果我们由某个特定的交叉路口的上道开始，并且沿曲线奔驰，到第一个交叉路口由下道过，到第二个交叉口由上道过，如此继续下去。到一定时候，我们将回到已经经过的交叉路口（A的二重点）。原来假设是在交叉路口的上道通过的，那么因为上下道交错变换，再次由上道通过这个交叉路口的情形会发生吗？不会，因为按照定理，正好说明这种情形是不可能出现的。根据定理，我们在回到原来的交叉路口以前，必须偶数次通过A的二重点或交叉路口。因为开始时是从上道通过最初的交叉路口，然后我们需要从下道通过第一个交叉路口，从上道经过第二个交叉路口，……从上道经过最后一个交叉路口（由于是偶数次）。因此，当回到原来的交叉路口时，将是由下道经过，而这正好是我们所希望的安排。二重点的新解释——空间曲线的交叉点，使我们得到这样的结果：如果我们从上面通过Q点，那么第二次一定是从下面通过它。

图 32 和 33 所示的空间曲线在平面上的投影，就是图26和图 27 所绘的曲线。每一个二重点可以看成是空间曲线的交叉点。这些空间曲线是打"结"的。沿曲线的投影行进时，会上下交错变换的那种结，叫作"交错结"。严格地说，图 32 中的结不是真正的结，因为对于这个具有二个小圈的曲线，可以把小圈扭放开为没有结的圆。而图 33 表示真正的结，因为不割断它就不可能变为圆。

由图 34 看出，至少在没有按某种方式变形前，并不是所有的结都是交错结。存在着非交错结的这个事实，也许最清楚地表明，我们的定理（不论是二种形式的哪一种）并不是无足轻重的。



图 32

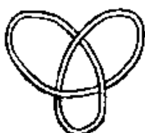


图 33

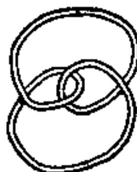


图 34

[注] 虽然在材料上本篇和第二篇有很多联系，但所用的概念在本质上是不同的，读者应很好地认识到二篇的相互独立性。在第二篇给出的是曲线网，并且讨论了它的行走路径。这里给出的是一个"闭路"，不存在它如何行走的问题。若用第二篇的语言，那么现在的曲线只有四阶连结点，但现在叫二重点。最后，在二重点，不能随便结合曲线的各个部分；会合于二重点的四个部分，因为曲线在这里本身是相交的，所以只能把对着的二部分结合起来，使成为二对。

[注 2] 是次数，而不是二重点的点数。在图 26 里的次序 1221 中，在 1 和 1 之间，只有单独的点 2，但我们二次通过了它。

11. 数的素因子分解是唯一的吗？

1. 我们从一个可以分解因子，并且一直可以分解到最后只是素因子的已知数着手。例如 60 可分解为 $6\cdot10$，6 又是 $2\cdot3$，且 10 是 $2\cdot5$，最后是

$$60=2\cdot3\cdot2\cdot5.$$

并且这四个因子都是素数。

还用这个例子，我们首先把 60 分解为

$$60=4\cdot15,\ 4=2\cdot2,\ 15=3\cdot5,$$

由此我们有

$$60=2\cdot 2\cdot 3\cdot 5.$$

在两个分解式中出现的因子是相同的素数，并且每个素数的次数也是相同的。

如果按照大小次序写出素数，那么两个分解式都是

$$60=2^2\cdot 3\cdot 5.$$

这里，二种分解式最后有相同的结果好象是明显的，因为我们经常遇到的就是这种情形。在算术中我们已经知道，如果一个数可以一直分解下去，直至最后都成为素数因子，那么不管怎样分解它，我们总是得到相同的因子。

这个命题是正确的，但它真是能明显成立的吗？且来考虑一个较大的数吧。分解 30031 要作大量的运算。经多次试验后，发现它可分解为 $59\cdot 509$，并且这些因子是素数。现在，根据什么，我们可以毫不含糊地说，不论再怎样试验也不能得到另一个完全不同的分解呢？

这个问题的提出就是违反我们的直觉概念的，我们的直觉认为，已知数必定有完全确定的素因子。后面第 2、3 节的目的就是要说明这些直觉概念不是真正的依据，因而说明这确实是一个问题。本篇其余部分将证明数的素因子分解是唯一的。

2. 为了摆脱我们所习惯的想法，将介绍一种不为大家所熟悉的数的系统。这是一些形如 $a+b\sqrt{6}$ 的数，其中 a 和 b 是普通的整数。例如 $12+5\sqrt{6}$，$\sqrt{6}-2$，和 $3\sqrt{6}$ 是这样的数，而 $2+3\sqrt{16}$ 则不是。普通整数不排除在外，事实上它们只是 $b=0$。它们形成我们这个数系中的一部分，因此，这个新的数系是普通整数集的一个扩充。

新数之间所实施的运算，可以很自然地得到。这个过程

是代数中所熟知的。两个这样的数的加、减，通过举例

$$(3+\sqrt{6})+(5+\sqrt{6})=8+2\sqrt{6}$$

就可以明白。按照代数的通常法则进行的下列乘法，将会表明任意二个数如何相乘：

$$(3+\sqrt{6})(3-\sqrt{6})=9-6=3,$$

$$(\sqrt{6}+2)(\sqrt{6}-2)=6-4=2,$$

$$(3+\sqrt{6})(\sqrt{6}-2)=3\sqrt{6}-6+6-2\sqrt{6}=\sqrt{6},$$

$$(3-\sqrt{6})(\sqrt{6}+2)=\sqrt{6},$$

$$(3+\sqrt{6})(2+\sqrt{6})=12+5\sqrt{6},$$

$$(3-\sqrt{6})(\sqrt{6}-2)=-12+5\sqrt{6}.$$

至于除法就不需要再讲什么了。和在普通整数中一样，有时一个数能整除一个数，有时则不能。

在我们的数系中，6 也可以分解为$\sqrt{6}\sqrt{6}$，同时 6 也可按普通方法分解，

$$6=2\cdot 3=\sqrt{6}\cdot\sqrt{6}. \tag{1}$$

这里，表面上 6 可以用二种不同的方法分解。这就使我们想到 30031 是否有不同于 59·509 的因子分解的问题。表面上，(1) 提出的是类似的问题。

然而，这种情形可以用一个很自然的方法加以澄清。数 2 和 3 是普通素数。在普通数系中它们不能再分解，但在这个新数系中却是可以分解的。事实上，由上述的四个乘法例子中，我们有

$$2=(\sqrt{6}+2)(\sqrt{6}-2),\ 3=(3+\sqrt{6})(3-\sqrt{6}).$$

因此，分解式 $6=2\cdot 3$ 还可继续分解，我们有

$$6=(\sqrt{6}+2)(\sqrt{6}-2)(3+\sqrt{6})(3-\sqrt{6}). \tag{2}$$

因此，(1) 中的二个不同的因子分解，就是 (2) 中因子的不

同的配对结合。其中第一个就是最前面二个因子的结合以及最后二个因子的结合。第二个分解式是第一个和第四个因子的结合以及第二个和第三个因子的结合。显然还可能有其它另外的结合。例如，如果我们结合第1个和第3个因子，以及第2个和第4个因子，则有

$$6=(12+5\sqrt{6})(-12+5\sqrt{6}),$$

并且这个分解可以通过直接相乘进行验证。

这种情形和我们所熟知的情形并无不同。在澄清的过程中，我们并不一定要知道(2)中的因子是素数。所谓素数，我们现在显然是指在我们的数系中不能再分解的数。实际上，证明这四个因子是素数并不困难。

3. 现在我们转向另一个数系。这个数系是形如 $a+b\sqrt{-6}$ 这种数的集合，其中 a 和 b 仍是普通的整数。在这种情形下，我们仍然求象(1)那样的式子，但这时已不可能再用如(2)所作的那样来解释它。在这个数系中，数的运算如同在数系 $a+b\sqrt{6}$ 中一样容易。利用的代数法则和以前一样。对应(1)，我们现在有

$$6=2\cdot3=-\sqrt{-6}\cdot\sqrt{-6}. \tag{3}$$

类比于另一种情形，我们试分解2、3以及 $\sqrt{-6}$。然而，此时它们全是这个数系中的素数，我们不能再分解它们了。

在以后的讨论中，引进数的"正则"的概念是很方便的。数 $a+b\sqrt{-6}$ 的正则是这个数和 $a-b\sqrt{-6}$ 的乘积，

$$N(a+b\sqrt{-6})=(a+b\sqrt{-6})(a-b\sqrt{-6})=a^2+6b^2.$$

换句话说，在我们的系统中，一个数的正则是这个数和用 $-\sqrt{-6}$ 代替 $\sqrt{-6}$ 所得到的数的乘积。一个数的正则总是一个正的普通的整数。还有，二个数的乘积的正则，等于它们

的正则的乘积。按照定义，我们有

$$N(a+b\sqrt{-6})(c+d\sqrt{-6})=[(a+b\sqrt{-6})(c+d(\sqrt{-6})][(a-b\sqrt{-6})(c-d\sqrt{-6})],$$

而这四个因子可重新排列次序，给出

$$(a+b\sqrt{-6})(a-b\sqrt{-6})(c+d\sqrt{-6})(c-d\sqrt{-6}).$$

比较这些项，按照定义，正好是

$$N(a+b\sqrt{-6})\,N(c+d\sqrt{-6}).$$

如果 2 在我们的数系中可以分解为二个因子，应当有

$$2=(a+b\sqrt{-6})(c+d\sqrt{-6}),$$

并且

$$N(2)=N(a+b\sqrt{-6})\,N(c+d\sqrt{-6}).$$

但 2 的正则是$N(2)=(2+0\sqrt{-6})(2-0\sqrt{-6})=2\cdot2=4$,所以应有

$$4=(a^2+6b^2)(c^2+6d^2).$$

即 4 应分解为二个普通整数的乘积,其中每个数都形如 x^2+6y^2。利用普通整数把 4 分解的方式只有二种。或者二个因子都是 2，或者 1 个因子是 4，另一个是 1。在这里，二种情形都无济于事，这是因为 2 不能表示为 x^2+6y^2，并且 1 只能在$x=1$，$y=0$时才具有这种形式。因此，在这个数系中，2 要分解为二个因子，就必须有一个因子是 $1+0\sqrt{-6}=1$。我们不能认为此数系中的这个分解比普通数系中的分解 $5=1\cdot5$ 多了什么内容。因此 2 是我们这个数系中的素数。

按照完全相同的方法，可以认为 3 和$\sqrt{-6}$ 也是这个数系中的素数。代替正则 4，是正则 9 和 6，它们也应分解为形如 x^2+6y^2 的因子。

我们现在已经证明了,(3)是在这个数系中的数 6 的二个

不同的素因子分解。由于在这个数系中数的二个不同的因子分解已经出现，因而在普通数系中数的素因子分解是唯一的就不是明显的了。如果有某种在普通数系中数的因子分解是唯一的明显成立的理由，那么按照同样的理由，也可以断定在任何数系中也是明显成立的。但我们现在已经找到了一个数系，在这个数系中，因子分解的唯一性不是明显的，甚至还是不正确的。我们将看到，在普通数系中因子分解是唯一的，并且这是这个数系的特殊性质。在以后的证明中，我们将利用这个特性。

值得注意的是，希腊数学家已认识到这个问题，并且感到对于逻辑的完整性来说，必须加以证明，而又必须不用我们那种例子来澄清前面提到的表面情形。素因子分解的唯一性定理，欧几里德已经证明了，但表述有些不同，没有应用近代理论。我们将作的证明不同于欧几里德的证法，而且并不只是定理形式上的差别。

4．30 是 3 的倍数。它也是 5 的倍数。这个事实可以用 30 是 3 和 5 的"公倍数"的说法表示。一般说，二个数的公倍数同时是每个数的倍数。不管这二个数是什么数，它们必定至少有一个公倍数，因为如果数是 a 和 b，它们的乘积 ab 必是公倍数。对于 3 和 5，乘积 $3\cdot5=15$ 和 30 一样，是它们的公倍数。数 30 是公倍数是因为它是乘积 15 的二倍。同样的理由，显然 ab 的每一个倍数都是 a 和 b 的公倍数。因此，二个数总有无限多个公倍数。

乘积 ab 和它所有的倍数合在一起，不一定给出了 a 和 b 的全部公倍数。例如，$10\cdot15$ 的倍数是 150，300，450，…，而 10 和 15 的公倍数是 30，60，90，120，150，180，…。

显然，30 是 10 和 15 的公倍数中最小的一个。任意二个数 a 和 b 总有一个最小公倍数，这只需要试验由 1 到 ab 间的公倍数的数。任意二个数 a，b 至少有一个公倍数，因为 ab 本身就是一个。所有公倍数中至少有一最小的，它就叫作最小公倍数。我们首先证明：

引理 1. 二个数的任一公倍数是它们的最小公倍数的倍数。对于 $a=10$，$b=15$，这个引理说明 30 的倍数，即 30，60，90，$\cdots$，是 10 和 15 的所有公倍数。对于这种情形以及其它任一特殊情形，引理都是很容易验证的。然而，我们必须证明，命题对于一般情形都是正确的。

这个证明有赖于这样一个事实，即 a 和 b 的二个公倍数之差仍是 a 和 b 的公倍数。a的 2 个倍数的差仍是 a 的倍数，这在第一篇中已经讨论和利用过。对于 b，这同样是正确的。因此，如果二个数同时都是 a 和 b 的公倍数，那么它们的差可以被 a 或 b 整除，因而是 a 和 b 的公倍数。

设 m 是 a 和 b 的最小公倍数，并且令N是任一公倍数。那么，根据我们刚才看到的，$N-m$ 也是 a 和 b 的公倍数。如果我们再减 m，那么 $N-2m$ 仍是 a 和 b 的公倍数。不断的减去 m，可见

$$N-m,\ N-2m,\ N-3m,\cdots,$$

都是 a 和 b 的公倍数。因为 m 是所有公倍数中最小的一个，第一数 $N-m$ 一定不是负的。这对于以后的各数同样应当是对的，但最终这些数中必将成为负的。假设 $N-xm$ 是正数中的最后一个。它是 a 和 b 的公倍数，并且由于减去 m 得到的下一个数不再是正数，因而它不大于 m。因为 m 是最小公倍数，那么只有 $N-xm=m$ 这一种可能。因此，N

$=xm+m=(x+1)m$ 是 m 的倍数。

5．我们还可以讨论一下二个数的“公约数”。如果一个数 c 可以同时整除 a 和 b，就叫它为 a 和 b 的公约数。乘积 ab 是 a 和 b 的公倍数，根据引理 1，它是最小公倍数 m 的倍数。我们现在证明

引理 2．二个数 a 和 b 的乘积除以它们的最小公倍数所得的商，即，数 $d=\frac{ab}{m}$
必是 a 和 b 的公约数。

由关于 d 的方程，我们有

$$d\frac{m}{a}=b,$$

而且因为 m 是 a 的倍数，$\frac{m}{a}$ 实际上是一个整数。因而 b 是 d 的倍数，或者换句话说，d 是 b 的约数。完全按同样的方法，d 也是 a 的约数，因此 d 是 a 和 b 的公约数。

6．我们现在可以证明一个定理，根据它可以立即推出素因子分解唯一性定理。我们证明

定理：如果一个素数 p 能整除二个数 x 和 y 的乘积，那么 p 能整除 x 或 y，即它至少能整除一个因子。

我们设 p 和 x 的最小公倍数为 m。一方面根据假设乘积 xy 是 p 的倍数，并且它显然是 x 的倍数。因而它是 p 和 x 的倍数，由引理 1，它也是 m 的倍数。

$$xy=hm. \tag{4}$$

另一方面，由引理 2，数

$$d=\frac{px}{m} \tag{5}$$

是一个整数，并且是 p 和 x 的公约数。素数 p 的约数只能是 1 或 p。因而或者 $d=p$ 或者 $d=1$。在第一种情形，$d=p$ 是 x 的约数，即 p 能整除 xy 中的第一个因子。在第二种情形 $d=1$，那么(5)成为 $1=\frac{px}{m}$ 或 $m=px$。代入 (4) 中，$xy=h(px)$。我们消去因子 x，得到 $y=hp$。由此看出，p 能整除 xy 的第二个因子。不论哪种情形，p 至少能整除一个因子。

由此我们推论：**如果素数能整除若干个数的乘积，那么它至少能整除其中一个因子。**

因为它如果能整除，例如 xyz，那么它或能整除 x 或能整除 yz。如果它能整除后者，那么它整除 y 或 z。在任何情形下，它必然至少整除一个因子。

7．素因子分解的唯一性定理随之立即可得。如果 N 有二个素因子分解

$$N=p\cdot q\cdot r\cdot s\cdots=P\cdot Q\cdot R\cdot S\cdots.$$

那么 p 能整除 N，因而 P 能整除右端乘积。由推论，它整除右端的一个素因子。但如果一个素数整除另一个素数，那么根据素数定义，这二个素数必须相等。因而 P 必然在右端的素因子中出现。按同样的方法，q 以及左端的其它素数也必然在右端出现。因为左右端是可以相互交换的，所以右端所有素因子也必在左端出现。换句话说，二个因子分解所含的恰是相同的素数。

现在我们只需证明每一素数在两端出现的次数也相同。如果 P 在左端出现 a 次，而在右端出现 A 次，我们有

$$N=p^a q^b r^c\cdots=p^A q^B r^C\cdots,$$

如果 a 和 A 是不同的，它们中的一个(设为 A)较大。那么我们可以除以 p^a，得到

$$M=\frac{N}{p^a}=q^b r^c \cdots = p^{A-a} q^B r^C \cdots.$$

这表示了M的二个因子分解，同时 P 只出现在右端，而左端却没有。我们刚才已经证明，在任何一数的二个因子分解中，相同的素数必然出现在二个因子分解中。在M这样的特殊情形下，也应当是正确的，因而 a 和 A 不可能是不同的。因此，我们有 $a=A$，并且用同样的方法，$b=B$，$c=C$，……。即每一个素数在每一个因子分解中出现的次数相同。

自然我们会感到奇怪，为什么同样的证明对于第 3 节中的数系 $a+b\sqrt{-6}$ 就不能成立。事实上，几乎所有的证明在这个数系中都能进行。只是一小部分，就是引理 1 不能成立。因此，这个引理是我们证明中的实质性的一步。

12. 四色问题及五色定理的证明

1．1879 年凯利讨论了如下的问题。一张地图为了区分不同的国家，常印上各种颜色。最好是每一个国家都印上一种不同的颜色，但这需要的颜色种类太多了。习惯上的代替办法是，只要注意彼此挨着的国家用不同的颜色就可以了，从而可能只需用少数几种颜色。图 35 a 是表示一个孤岛的地图，只需要三种颜色。海用蓝色，二个国家用另二种颜

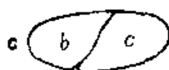


图 35 *a*

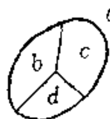


图 35 *b*

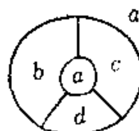


图 35 *c*

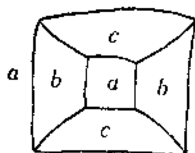


图 36

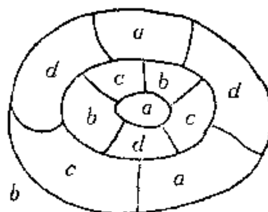


图 37

色。图 35 *b* 需要四种颜色。三个国家都与海相接，所以都不能和海取同一色，又因为每个国家与另外二个相邻，它们就需要三种颜色，总共着了四色。图 35 *c* 表明，即使我们不算海的颜色，也需要四色。这里岛的内部的国家与图 35 *b* 的海起着同样的作用。图 36 再次只需要三色，而较复杂的图 37 需用四种颜色。自然会想到越是复杂的地图，需要的颜色越多。

我们画过很多地图，但没有注意到，不管多么复杂的地图，上面的颜色往往不超过四种。另一方面，又总没有人能证明，对于可以设想到的地图，只要四种颜色就足够了。这个问题，没有任何数学知识的人都可能提出并完全理解，但直到现在尚未解决[1]。

然而，已经证明，任何地图都可用五种颜色着色。这里

可以着色的意思，是指没有二个有共同边界的国家是同一种颜色，但是只在角顶相交的国家可以有同一种 颜 色。（如跳棋盘的着色）而且，作为国家，我们是指连成一片的陆地，而不是由几个分离部分组成的政治单位。

本章的主要目的在于证明，每张地图都可用五种颜色着色。在证明中，假定地图所表示的是孤岛。若每个孤岛和海能用五种颜色着色，则由几个岛组成的地图，用到的也恰好是在每个岛中同样的颜色。

2．作为第一个预备知识，我们将证明 欧拉定理。这个定理涉及到任意一张地图的顶点 v（角顶），面 f（国家）和边 e（边界）的个数。这是一个具有一般性的定理，并且有很重要的多方面的应用，欧拉发现的这个定理，实际上笛卡尔早已了解。这个定理断定

$$v+f=e+2。\tag{1}$$

例如，在图 36 有 8 个顶点（即至少有 三 个 国 家相交的点），6 个面和 12 条边（每一条边 都从一个顶点伸展到下一个顶点）。事实上，我们有

$$8+6=12+2。$$

为了证明这个定理，我们暂时离开有关地图的设想，而是用堤和田所组成的系统来表示图形。我们用田表示面，边现在就是把田分开的堤，而整个外部面积则是湖水（原来地图上的海）。我们现在想依次拆毁一些堤，使 所有的田都灌上湖水（图 38）。显然，这没有必要把所有的堤拆毁，因为二边都已有水的堤就可以保留。如果我们只拆毁仅一边有湖水的堤，那么每拆一个堤就多灌了一块田。外部区域（相应于地图中的海）原来就有水，那么还有 $f-1$ 块田需要灌水。

因为灌水的过程是要使所有的田都被灌上水为止，所以最后将拆毁的恰好是 $f-1$ 个堤。

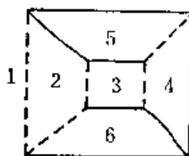


图 38

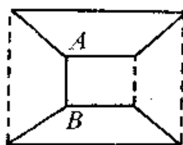


图 39

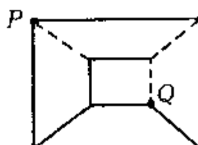


图 40

我们现在研究的堤的系统原来是左边的样子。在拆堤时，还要满足下列两个要求。

Ⅰ．在不能涉水的条件下一个人能沿堤岸从任一个顶点走到任一其他顶点。在所有的堤未拆毁前，这一定能作到，因为一开始就假定地图表示的是一个孤岛。在灌田的过程中，有可能出现 AB 这样的堤(图 39)，当把它拆毁时，就把整个系统分割为二个完全分离的岛，那么这二个分离的系统完全被湖水包围，因而从 A 不能走到 B。那么显然，这样的堤将不能拆毁。即二边都有水的堤不能拆。

Ⅱ．一个送信者从任一顶点到任一其他顶点只可能有一条途径。否则，若从 P 到 Q 有二条不同途径，则这二条途径将包围一个面积(图 40)。于是，未拆的堤所组成的二条道路将围住一块旱地，这与所有的田都要被灌水矛盾。

设开始时 P 固定，且到每一点都只有一条路。这条路的最后一条边必然在要到达的顶点的前边。所以这条边由这个点完全确定。这就建立起边与顶点的对应关系，即每条边和它的终点对应。这样，有多少条未拆的堤就有多少个顶点。

而又因开始的P不是终点，所以未拆的堤的数是$v-1$。在所有的堤中，有$f-1$个拆毁和$v-1$个未拆毁，所以堤的总数是

$$e=(f-1)+(v-1)。$$

如果我们去掉括号并把数字移项，就立刻得到欧拉定理。

3．证明五色问题的最后一个预备知识是，只需要对在任一顶点相交的国家都不多于三个的地图作证明。若在某一个点相交的国家超过三个，如图41a，则可以另作一新地图，它除了围绕顶点的小国家外和原图一样。这个新地图比

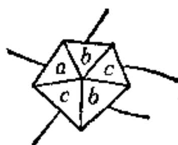


图 41 a

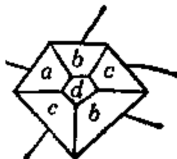


图 41 b

原图多了一个国家和若干顶点，但在顶点只有三个国家相交，多于三个国家相交的点被去掉了。如果对于每一个相交多于三国的顶点都这样作，那么最后就有一张没有任何这种点的新图。现在如果我们证明五色对于这种新图是足够了，那么，五色对于原来的地图也是足够了。一种可能的着色法可从图 41 b 的字母上看出来。如果在新图上的每个国家已着了某些颜色，那么原图上的色与去掉增加的那个小国后一样(图 41 a)。这里没有违反顶点相交的二个国家沿边界不是同一色的原则。

任一顶点至少有三个国家相交，但是我们又已知道，只需研究在顶点相交不多于三个国家的地图。因而我们只需研究在每个顶点恰恰只有三个国家相交的地图。

4．现在可以转向证明本身。我们要研究每一个国家的边界上的交点数。如果一个国家的边界上只有一个交点或没有交点，那么它只有一个邻国，所以只要给它上一种与邻国不一样的颜色即可。这当然很容易，所以将略去它们，并假定在今后的证明中不再出现。

设 f_2 是边界上有二个顶点的国家数，f_3 是有3个顶点的国家数，……，那么所有国家数等于这类国家数的总和，即

$$f=f_2+f_3+f_4\cdots \tag{2}$$

f_2 这类国家有二个顶点，所以有二条边界，总共的边界数为 $2f_2$，f_3 有三个顶点，则每个国家有三个边界，共有边界 $3f_3$，如此等等。最后计算总边界数，但每一个边界都计算了二次，因为它在二个国家分别计算了一次。所以

$$2e=2f_2+3f_3+4f_4+\cdots \tag{3}$$

按同样办法计算顶点数，但恰好每个顶点有三个国家相交，所以

$$3v=2f_2+3f_3+4f_4+\cdots \tag{4}$$

从(3)与(4)得到

$$3v=2e。 \tag{5}$$

欧拉公式(1)乘6，得到

$$6v+6f=6e+12,$$

由于(5)，上式右端为

$$9v+12$$

因而　　　$6f=3v+12$

或，由(2)和(4)得

$$6(f_2+f_3+f_4\cdots)=(2f_2+3f_3+4f_4+\cdots)+12。$$

化简，有

$$4f_2+3f_3+2f_4+f_5=12+f_7+2f_8+\cdots \qquad (6)$$

由此可见，每张交点有三个国家相遇的地图，至少肯定有一个国家其边界上的顶点数是少于6的。因为如果没有这样的国家，也就没有二个顶点的国家。因而 $f_2=0$。同样，$f_3=f_4=f_5=0$。因此(6)的左边是0，而右边至少是12。如图 35_a, 35_b, 36 和 37 的地图，没有一个国家其边界上的顶点数多于6个，于是在所有这些情形下，(6)的右端都是12。另一方面，我们有

在图 35_a　$f_2=3$，$f_3=0$，$f_4=0$，$f_5=0$，

35_b　$f_2=0$，$f_3=4$，$f_4=0$，$f_5=0$，

36　$f_2=0$，$f_3=0$，$f_4=6$，$f_5=0$，

37　$f_2=0$，$f_3=0$，$f_4=0$，$f_5=12$.

这些例子中的每一个都只有一个 f_i 不等于零。图 35_c 的情况则不同，因为 $f_2=0$, $f_3=2$, $f_4=3$，其它的仍是零。

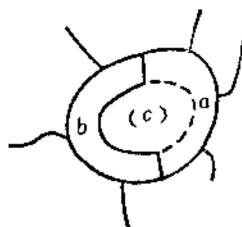


图 42

现在知道，我们的地图至少有一个国家其顶点数少于6个。那么，就是2，3，4或5个顶点。现在分别讨论这四种情形。

Ⅰ. 有一个有二个顶点的国家。这个国家恰好有二个邻国。设想二个边界中去掉一个（图42中的虚线）。则新图用 f-1 个国家代替了原来的 f 个国家。假定有较少国家的新

图可着五色，则原图中取消边界的邻国用 a 色，另一个邻国用 b 色。如果我们放回边界，则原来有二个顶点的国家可以着 c 色，因为这个国家只有 a 色和 b 色的邻国，从而使 c，d 和 e 色都可用上。因而 f-1 个国家的新图若可用五色着色，则原图也一定一样可用五色着色。

Ⅱ. 有一个有 3 个顶点的国家。设 L 是有邻国 L_1，L_2，L_3的国家（图 43）。去掉L的一个边界，并假定有 f-1 个国家的新图不需要多于五色，则再加上边界后，只需要给 L 一个与 L_1、L_2、L_3 不同的色。L 可利用另外两种颜色中的任何一色。

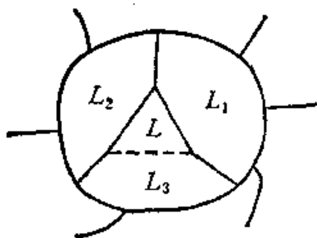


图 43

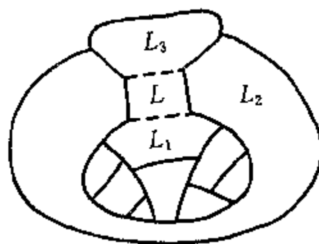


图 44

Ⅲ. 有一个有四个顶点的国家。用类似的方法，我们看到这里五色中至少有一种颜色可以利用，因为四个邻国可能用四种不同的色。但在这种情况下，一种新的困难出现了。图 44 中，L_2这种国家与L可以沿二个不同的边界相交，若去掉其中一个边界，则另一个也必须去掉，因为一个国家本身不能再用边界分开。因而出现了这样一种国家，其边界是由二个完全分离的边所组成，象一个环。但我们证明欧拉定理时未加说明地排除了这种国家。而且，我们可以避免形成

这种环形国家。若 L 与 L_2 一起组成一个环，则 L 的另二个边界必然属于由环形隔开的二个国家 L_1 和 L_3。因而 L_1 和 L_3 是不同的，而且是没有共同边界的国家，所以它们可用同一种颜色。因此去掉把 L 同 L_1 和 L_3 隔开的边界后，得到的新图有 f-2 个国家。如果这个少于 f 个国家的新图可着五色，则原图中 L 的三个邻国中有二个同一色。因而 L 可利用剩下三色中的任一色。

Ⅳ. 有一个有五个顶点的国家。同样的困难将会在一种更为复杂的形式中产生。不论是一个国家和 L 有二个不同的共同边界（图 45_a）或 L 的二个邻国可以在另一不同的边界相交（图 45_b 的 L_2 和 L_4），二种情形 L 都有二个邻国 L_1 和 L_3。它们是不相邻的不同的二国。因为很清楚，它们不能

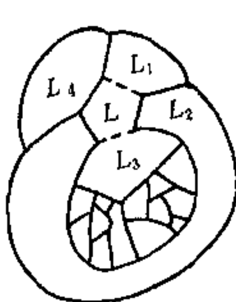


图 45 a

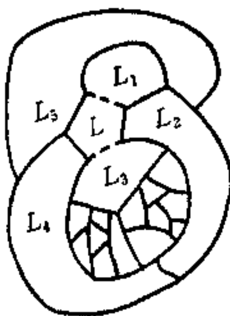


图 45 b

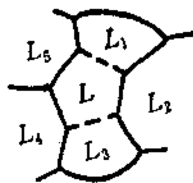


图 46

通过图 45_a 中 L 和 L_2 组成的环形或图 45_b 中 L、L_2 和 L_4 组成的环形。在这二种情形以及在不出现环形的简单情形下，L 有二个邻国 L_1 和 L_3 不在一个边界上相交（图 46）。现在去掉 L 分别与 L_1 和 L_3 相交的边界，则新图有 f-2 个国家，比原图少。如果再次假定新图能着五色，且 $L_1+L_2+L_3$ 有

色 a，L_2 为 b，L_4 为 c，L_5 为 d，共 4 色，那么再加上二个边界，L 还可用第五种色。

每张地图都将属于这四种情形之一，于是我们的简化程序就完成了。每一种地图都可由于去掉一个边或二个边而简化，简化后的图，国家减少，并且如可由五色着色，则原图也可。重复这个简化过程，直到不多于五个国家。那么不多于五个国家的地图显然可用五色着色，因而对于简化过程中的每张图也可以，最后对于原图也是对的。

5．我们已证明了平面上所绘地图的五色定理。然而对于球面上的地图，也可以利用同样的证明。国家的外面，通常到处是海，它遍布球面的剩余部分。我们以上的证明，在球面上逐步都能进行。这里就不再重复了。这个证明没有利用任何线段或角的相等性。它不包含任何全等性定理，或任何别的不能直接推广到球面上的概念，只利用了有关点、曲线和面积的相对位置的概念，而这些概念在球面上和在平面上是一样的。

如果我们研究象轮胎形那样的曲面上的地图，情形就不同了。在这样的曲面上画一张七个国家组成的地图，其中每个国家都与其它六国有共同边界是可能的。因而这个地图需要七色。除非有模型，在这种曲面上的地图是不容易直观化的，所以我们只满足于叙述这样的事实。即五色定理的证明不能推广到这种曲面。为什么我们的证明容易推广到球面，但推广到轮胎形曲面时就错了呢？

我们的证明在通常称之为圆环的轮胎形曲面上时，有二点是不能成立的。首先，在欧拉定理证明的最后，参见图 40，我们说过从 P 到 Q 若有二个途径它们就可包围一个面

积。第二点是，在这个证明的情形Ⅲ和Ⅳ时，我们曾说 L_1 和 L_3(图 44,45 a, 45 b) 不能相交，是由于它们被 L 和 L_2 或 L，L_2 和 L_4 组成的带环分割开了。在这二种情形下，平面上或球面上的证明依赖于这一事实，即就闭曲线一侧的点到达另一侧的点，若不横穿曲线是不可能的。但以轮胎形曲面而言，这是不对的。为了帮助我们理解，我们可直观想象这一曲面。让我们设想一下附有带环的土星（图 47)。这带环实际是由颇为松散的质点组成的立体图形。这个带环是一个

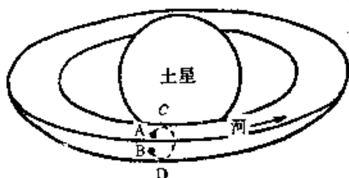


图 47

轮胎。设有一河绕带环流过，而这个带环的一侧和土星还有一段距离。若我们从河岸一侧的 A 开始，我们能不能不横渡该河而走到对岸的 B 点呢？能！只要我们离开河直接走到点 C，越过离土星最近的这一边到达 D，然后在环下来到 B。

这些情形说明，当依赖于直观时要多么小心，直观是怎样地容易使人误入歧途。图 40 有关的那部分证明，当时直观上似乎很明显，但现在我们知道实际上需要一个逻辑证明，上述直观证明必须依赖平面和球面的特殊性质，而它们对于轮胎形曲面却不一定仍然成立。

我们将不涉及平面和球面性质的深刻分析。但我们所假定的那些性质是可以证明的，虽然这并不很容易。在轮胎形

曲面的情形下，可以证明必须用公式

$$v+f=e$$

来代替欧拉公式。于是五色定理的证明也可以推广到轮胎上，即在这种曲面上的地图上色，七种颜色就足够了。这就产生了一件很有意思的事情：在较为复杂的轮胎形曲面上，上色问题完全能加以解决，而在较为简单的平面和球面上，还不知道究竟需要五色还是四色。

［注］ 1977年12月9日《人民日报》钱玉森的文章“现代科学技术”中写道：“去年数学界哄动一时的一件事，是用电子计算机证明了数学方面的四色定理。要证明四色定理很难，数学家经过上百年的努力，证明不了。去年美国数学家用电子计算机证明了。他们看到这个问题要证明并不是不可能，而是证明的步骤、程序很复杂，人一辈子的时间也证不完。他们就把程序编好，交给高速的电子计算机去干。高速电子计算机也用了一千多小时才证出来。”

——译注

13. 正多面体

1. 我们将要利用欧拉定理，求得一种完全不同类型的结果。我们将研究的问题是：存在正多面体吗？若存在，究竟有多少种？一个多面体是由被称为面的一些平面的部分所围成的立体图形。按照欧几里德定义，所谓“正”多面体，是指其所有面都是等角和等边的全等多边形（正多边形）的多面体。

如果我们应用更一般的定义，则我们的问题将会更广泛，答案会更满意。我们称一个多面体是“正”的，是指它的

所有面的边数都相同，而且在每一顶点(角顶)有相同数量的面交会在一起。在这个定义中，我们没有涉及等边、等角或面积，也没有其它与大小有关的量。在此定义中，只用到某些部件的数目。

我们令 φ 表示在每一面上的顶点数。如果多面体为三角形组成，那么 $\varphi=3$，等等。在每一顶点相交的面的个数，可表示为 ε。

因为多面体每个面至少有三个顶点。我们有

$$\varphi \geqslant 3。\tag{1}$$

在立体图形中，一个顶点至少有三个面相交，这样我们有

$$\varepsilon \geqslant 3。\tag{2}$$

我们令 v、f 和 e 分别表示多面体的顶点、面和棱数。

我们可以把多面体想象成是由弹性物质如橡皮组成的中空的东西，因而可把它充气成一个球，则多面体的面就成了这个球面的一部分，棱就成为球面曲线的一部分。如果把这个球看作地球仪，则原来的多面体可以看成表示 f 个国家的地图。每个国家可由一个面组成。国家的边界由多面体的棱形成，则边界数必为 e。同样地，地图和原来的多面体一样，有相同的顶点数 v。每个国家有相同的顶点数和边界数。因而在数 $f_2, f_3, \cdots$(上篇第 4 节)中不等于 0 的只有一个，而且这一个数必然等于面 f 的总数。在每一个顶点恰好有 ε 个国家相交。这种地图的例子见图 35b，36，37。在这些例子中，我们有

$$\varphi=3,\ \varepsilon=3;\ \varphi=4,\ \varepsilon=3;\ \varphi=5,\ \varepsilon=3.$$

由于可把多面体充气成地图，那么，就如同对待地图一

样，我们可以把欧拉定理应用于多面体。这就是说，任意多面体都有表示顶点数、面数和棱数之间的关系的公式

$$v+f=e+2。\tag{3}$$

根据历史，欧拉发现的公式就是这种形式，而且它通常被列成这种形式。

2．正多面体的面有 φ 个顶点和 φ 条棱，那么 f 个面就有 $f\varphi$ 条棱。但是，这里每条棱都计算了二次，因为每一条棱都是二个不同的面的棱。因而我们有

$$f\varphi=2e。\tag{4}$$

类似，在每个顶点有 ε 个面相交，因而有 ε 条棱相交。这样给出 $v\varepsilon$ 条棱。但每条棱也计算了二次，因为每条棱有两个端点，因而有

$$v\varepsilon=2e。\tag{5}$$

由(3)有

$$v+f-e=2,$$

乘以 2ε，就变为

$$2v\varepsilon+2f\varepsilon-2e\varepsilon=4\varepsilon。$$

现在，由(4)，我们可以用 $f\varphi$ 代替 $2e$，又由 (4) 和 (5) 知 $f\varphi=v\varepsilon$，所以可以用 $f\varphi$ 代替 $v\varepsilon$，从而有

$$2f\varphi+2f\varepsilon-f\varphi\varepsilon=4\varepsilon,$$

或

$$f(2\varphi+2\varepsilon-\varphi\varepsilon)=4\varepsilon.\tag{6}$$

因为 f，4ε 都是正数，那么括号内的因式也必然是正数，所以有

$$2\varphi+2\varepsilon-\varphi\varepsilon>0。\tag{7a}$$

公式(1)和(2)给出了 φ 和 ε 的下界；现在我们要从(7a)找到上界。首先我们变换(7a)的代数符号，得到

$$\varphi\varepsilon-2\varphi-2\varepsilon<0. \tag{7b}$$

如果我们把(7 b)的左端与乘积

$$(\varphi-2)(\varepsilon-2)=\varphi\varepsilon-2\varphi-2\varepsilon+4 \tag{8}$$

作比较，我们看到乘积只是加了一项 4，所以我们应在(7 b)的二边加 4

$$\varphi\varepsilon-2\varphi-2\varepsilon+4<4,$$

与(8)比较，得

$$(\varphi-2)(\varepsilon-2)<4。 \tag{9}$$

由(1)，(2)可见$(\varphi-2)$，$(\varepsilon-2)$至少是 1。因为$(\varphi-2)$和$(\varepsilon-2)$是正整数，且它们的乘积小于 4。所以求所有这类数的乘积并不困难。它们是：

$$1\cdot1,\ 1\cdot2,\ 2\cdot1,\ 1\cdot3,\ 3\cdot1, \tag{10}$$

总共是五个乘积。二个正整数的任一个其它乘积或等于 4 或是一个更大的数。

由此可见，乘积小于四的二个正整数只有五种情况，因而推导出正多面体只有五种。

3．在(10)的五个乘积中，设第一个因子是$(\varphi-2)$，第二个是$(\varepsilon-2)$。那么，我们可得到φ，ε的五对值：

φ	ε
3	3
3	4
4	3
3	5
5	3

我们已经直接由欧拉定理得到此表。现在，如果我们记得φ和ε的意义，我们从本表可以看到，正多面体如果存

在的话，只能由三角形、四边形或五边形作为面。而且，3，4 或 5 个面必相交于顶点。现在由(6)我们有

$$f=\frac{4\varepsilon}{2\varphi+2\varepsilon-\varphi\varepsilon},$$

利用(8)，这可写成

$$f=\frac{4\varepsilon}{4-(\varphi-2)(\varepsilon-2)}。\qquad (11)$$

由(4)得

$$e=\frac{f\varphi}{2}=\frac{2\varepsilon\varphi}{4-(\varphi-2)(\varepsilon-2)}。\qquad (12)$$

由(5)和(12)，我们求得

$$v=\frac{2e}{\varepsilon}=\frac{4\varphi}{4-(\varphi-2)(\varepsilon-2)}。\qquad (13)$$

最后这三个公式只给了我们相当于 φ 和 ε 五对可能的值的 f、e 和 v 的一个值。这些值可由下表列出：

φ	ε	f	e	v	
3	3	4	6	4	正 4 面体
3	4	8	12	6	正 8 面体
4	3	6	12	8	正 6 面体
3	5	20	30	12	正20面体
5	3	12	30	20	正12面体

因而只有五种正多面体。它们也称为五个柏拉图体，并如表所示，它们是按面的数目命名的。

4．这个表有一个很特殊的性质。如果 φ 和 ε 互换，则

20 面体和 12 面体互换，8 面体和 6 面体互换，但 4 面体没有改变。我们也可以不参考此表而从前面的公式知道这些。由于条件(1)(2)和(9)对于 φ 和 ε 是对称的，当 φ 和 ε 互换时，任何可得到的一对 φ 和 ε 的值仍可得到。并且(12)对于 φ 和 ε 也是对称的，所以它们互换时 e 不改变。最后，由(11)和(13)可知，φ 和 ε 的互换，只互换了 v 和 f。

这个关系，从纯几何的理由来看也是明显的。我们只需要选择多面体每一面的任意一点，并且把这些点看作是新多面体的顶点。新多面体的棱是由原多面体所有相邻二个面上的二个顶点连接而成。那么新多面体在原多面体的每一面上恰有一顶点，恰有一新棱横穿过每一旧棱的上空，恰好一个新面切去一个顶点。因此，二个多面体有同样的棱数 e，而 v 和 f 互换。

5．还有重要的一点没有讨论。到现在为止，我们只说明，如果所有的面有相同的顶点数，并且在每一顶点相交的面的数目相同，那么面、棱和顶点数必对应于表中的五种可能情况。因而**至多**有五种正多面体。

整个表包括了所有可能的正多面体，但我们并没有说明，这五种多面体实际是否可以作出。完全可以设想，存在着某些我们并没有考虑在内的更多的限制。这些进一步的限制，可以除去表中所列的一种或几种类型，总之，我们所讨论的是关于正多面体的**必要条件**而不是**充分条件**。

实际上，我们对球面上的“正”地图的讨论，比对正多面体本身的讨论更为重视。这些地图只是多面体充气为球面后的多面体。现在我们可以确切举出对应于表中每一类型的地图。图 35 b 表示四面体，图 36 是六面体，而图 37 是十二

面体。对应于八面体和二十面体的是图 48 和图 49。虽然所有这些图都是画在平面上的，但都可以变换到球面上。现在，如果我们不考虑多面体充气为球面的形变，那么我们看到，已建立的关于正多面体的必要条件也是充分的。

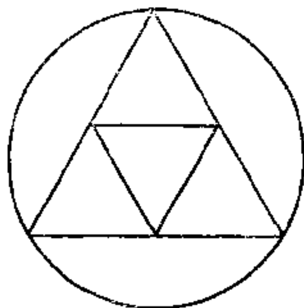


图 48

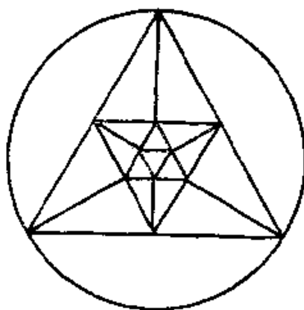


图 49

这样就完成了我们对问题的讨论。但是我们还没有证明，我们的图形，按照欧几里德狭窄的正多面体的定义是可以作出来的。(正多面体是其所有面都是全等多边形的多面体)要证明这点，需要运用类型完全不同的概念和理论。我们前面的研究，利用的是伸缩形变下仍不改变的性质。全等不是这样的性质。所以现在的证明需要利用度量几何。在度量几何中，长度和角的相等性很重要，利用度量几何可以证明上面的多面体按欧几里德定义是能作出的。我们不打算作此证明。狭义的恰好是五个正多面体的存在性，古时就有记载。这个证明是柏拉图的学生查达图斯的贡献，欧几里德在它的«几何原理»第 XIII 卷中末尾给出了它。

14. 毕达哥拉斯数和费马定理

1. 根据毕达哥拉斯定理[1]，直角三角形斜边上的正方形的面积，等于二腰上正方形面积的和。反之，如果在三条线段中，有一条的平方等于另外二条的平方的和，那么，这三条线段可以组成为一个直角三角形。方程 $a^2+b^2=c^2$ 表示了这样一个事实，即长为 a, b, c 的线段是直角三角形的边。

我们早在第四篇中已经知道，一个等腰直角三角形的腰和斜边是不可通约的，即没有整数 a 和 c 能满足方程 $2a^2=c^2$。是否存在某一个直角三角形，它的边是可通约的呢？换句话说，是否有满足方程

$$a^2+b^2=c^2 \tag{1}$$

的三个整数？一个简单而很熟悉的例子

$$3^2+4^2=5^2 \quad 或 \quad 9+16=25$$

说明回答是肯定的。

是否有另外的答案呢？我们怎样才能找到它们呢？本篇将求得这个问题的完整的解答。

2. 如果我们有(1)的一个解 a, b, c，我们用任一整数乘每一项 a, b 和 c，可以很容易求得另外一个解。例如 3，4 和 5 是解，我们用 2 乘，得 6，8，10。于是有

$$6^2+8^2=10^2.$$

如果 n 是任意正整数，那么一般的 $3n$，$4n$，$5n$ 就是解。同样的方法，如果 a, b, c 是任意解，那么 an，bn，cn 也是解，这是因为由 $a^2+b^2=c^2$，我们有 $a^2n^2+b^2n^2=c^2n^2$

或 $(an)^2+(bn)^2=(cn)^2$。这个寻求新解的方法是比较平常的，不那么吸引人。比较有趣的方法是寻求基础解，就是那些不可能象刚才那样由另一个解乘以某个整数而得的解。如果一个解 a, b 和 c 没有公因子，我们就称这样的解为"简化解"。因而，3，4，5 就是简化解。

如果两个或更多个数没有公因子，我们就说它们"互素"。在简化解中，数 a, b 和 c 中的任一对都互素。如果 a 和 b 有一个公因子设为 d，那么 d 的任一因子也是它们的公因子。一定有某个素数 p 能整除 d；至少当 p 是 d 本身时就是这样。那么 a 和 b 应有公因子 p，而且我们可以写成

$$a=pa_1, \quad b=pb_1.$$

则方程(1)成为

$$p^2(a_1^2+b_1^2)=c^2,$$

由此可见 p^2 应能整除 c^2。那么 p 应能整除 $c^2=c\cdot c$，并且因此能整除二个相同因子中的一个。即如同对 a 和 b 一样，p 能整除 c。根据同样的方法，a 和 c 或 b 和 c 的公共素因子是所有三个数的公因子。

3．现在我们来研究一下(1)的解 a, b 和 c，其中 a, b 和 c 的任意一对都是互素的。这些数中不可能有二个是偶数，即可以被 2 整除的数，至多一个是偶的。然而不可能所有三个数都是奇的。一个奇数 $a=(2l+1)$ 的平方是 $a^2=4l^2+4l+1$，仍是奇的。如果 a 和 b 是奇的，那么 a^2 和 b^2 是奇的，而 a^2+b^2 却是偶的，显然这不能等于奇数 c 的平方。

只剩下数 a, b, c 中 2 个是奇的，一个是偶的这种情形。进而我们可以看到 c 必定是奇的，因为如果 c 是偶的，它可以被 2 整除，而 c^2 可被 4 整除。另外二个数必是奇的，设

$$a=2l+1,\ b=2m+1,$$

并且我们求

$$a^2+b^2=(4l^2+4l+1)+(4m^2+4m+1)$$
$$=4(l^2+l+m^2+m)+2.$$

这个数是偶的，但被 4 除后还剩余数 2，因此它不能等于能被 4 整除的 c^2。

现在我们只可能有 c 是奇的，数 a 和 b 中一个是偶的，另一个是奇的情形。我们不妨令 a 是奇数，b 是一个偶数。例如，在我们的例子中 $a=3$，$b=4$，$c=5$.

4．等式(1)可以写成如下形式

$$b^2=c^2-a^2=(c+a)(c-a). \tag{2}$$

其中 $(c+a)$ 和 $(c-a)$ 是二个奇数的和与差，它们都是偶数，仅有的公因子是 2。换句话说，$\frac{c+a}{2}$和$\frac{c-a}{2}$是互素的，否则，我们可以假设 d 能整除这二个数。那么

$$\frac{c+a}{2}=df,\quad \frac{c-a}{2}=dg,$$

加或减这二个式子，得到

$$c=d(f+g),\ a=d(f-g)$$

那么 d 可同时整除 a 和 c，这与它们是互素的假设矛盾。

由于 b，$c+a$，$c-a$ 都是偶数，我们可以把 (2) 式写成如下形式

$$\left(\frac{b}{2}\right)^2=\frac{c+a}{2}\cdot\frac{c-a}{2}. \tag{3}$$

其中的分数都只是表面的，因为每个分数实际上都是一个整数。这个等式表示$\left(\frac{b}{2}\right)^2$是二个互素的因子$\frac{c+a}{2}$和$\frac{c-a}{2}$的

乘积。现在我们到了这个证明的关键的一步。我们证明$\frac{c+a}{2}$和$\frac{c-a}{2}$必定都是平方数。如果$\frac{b}{2}$可以分解为素因子的乘积，

$$\frac{b}{2}=p^{\alpha}q^{\beta}r^{\gamma}\cdots,$$

其中 p, q, $r\cdots$是不同的素数，那么我们有

$$\left(\frac{b}{2}\right)^2=p^{2\alpha}q^{2\beta}r^{2\gamma}\cdots.$$

$\left(\frac{b}{2}\right)^2$的所有素因子必然同时会出现在$\frac{c+a}{2}$和$\frac{c-a}{2}$的乘积中。然而，因为$\frac{c+a}{2}$和$\frac{c-a}{2}$没有公因子，所以每个素因子 p 必然或只出现在$\frac{c+a}{2}$中，或只出现在$\frac{c-a}{2}$中。因而$\left(\frac{b}{2}\right)^2$的素因子是按照每个素数的幂 $p^{2\alpha}, q^{2\beta}, r^{2\gamma}, \cdots$ 或在$\frac{c+a}{2}$或在$\frac{c-a}{2}$内的方式分布在$\frac{c+a}{2}$和$\frac{c-a}{2}$的乘积之中。因此，$\frac{c+a}{2}$和$\frac{c-a}{2}$中所含的只是素因子的偶数幂，因而每一个都是平方数。

5．我们现在可以写出

$$\frac{c+a}{2}=u^2,\quad \frac{c-a}{2}=v^2, \tag{4a}$$

$$\left(\frac{b}{2}\right)^2=u^2v^2, \tag{4b}$$

其中 u 和 v，如同 u^2 和 v^2 一样，是互素的。由（4b）我们有

$$b=2uv, \tag{5}$$

又通过加或减方程（4 a）的二式，我们求得

$$c=u^2+v^2,\ a=u^2-v^2。 \tag{6}$$

由于 c 和 a 同时是奇数，平方数 u^2 和 v^2 中的一个必是偶的，而另一个是奇的，否则它们的和或差将会是偶的。u 和 v 也必然一个是奇的，另一个是偶的。我们称二个这样的数是"互反数对"。

现在我们已经证明，如果 a, b, c 是（1）的简化解，那么，a，b 和 c 可以用二个互素的，且是互反数对的数，以(5)或（6）的形式表示出来。在我们原先的例子中，$a=3$，$b=4$，$c=5$，我们有

$$u^2=\frac{5+3}{2}=4,\quad v^2=\frac{5-3}{2}=1,$$

$$u=2,\ v=1,$$

并且和(5)一致，

$$b=4=2\cdot2\cdot1=2uv.$$

6．迄今为止，我们只是求得简化解的必要条件。我们由简化解 a, b, c 开始，并且确定了 u 和 v。为使讨论完整，我们必须证明条件是充分的，即当 u 和 v 互素，且是互反数对的数时，由(5)和(6)给出的 a, b, c 总是简化解。我们还应加上条件 $u>v$ 才能保证 a 是正的。

首先，由简单的计算得到

$$(u^2-v^2)^2+(2uv)^2=(u^2+v^2)^2,$$

所以由(5)和(6)给定的数满足方程(1)。第二，要说明我们得到了一个简化解，只需记得 u 和 v 是互素的，并且一个是偶的，一个是奇的。由此可见，方程(6)中的 a 和 c 同时是

奇的，因此 a,b,c 没有公因子 2。 a 和 c 也不能有一个奇的公因子。不然的话，若有一个公因子是奇的素数 P(不同于 2 的某个素数)。我们可以写为

$$c=Pc_1,\ a=Pa_1,$$

并且由(4 a)，应有

$$2u^2=c+a=P(c_1+a_1),$$

$$2v^2=c-a=P(c_1-a_1)。$$

这些方程说明 P 可同时整除 $2u^2$ 和 $2v^2$。因为 P 不同于 2，那么应整除 u^2 和 v^2，但这是和 u 与 v 是互素的事实不相容的。

下面列出由(5)和(6)导出的一些毕达哥拉斯数的例子：

$u=2,\ v=1;\ a=3,\ b=4,\ c=5;$

$u=3,\ v=2;\ a=5,\ b=12,\ c=13;$

$u=4,\ v=1;\ a=15,\ b=8,\ c=17;$

$u=4,\ v=3;\ a=7,\ b=24,\ c=25;$

$u=5,\ v=2;\ a=21,\ b=20,\ c=29;$

$u=5,\ v=4;\ a=9,\ b=40,\ c=41。$

显然，在此表中 b 不仅是偶数，同时它总是 4 的倍数。这是由于 $b=2uv$ 而 u 或 v 是偶数。

7. 我们已经完全解决了方程(1)所涉及的问题，现在的目的是作一系列的推广。

我们可以对方程

$$x^3+y^3=z^3, \tag{7 a}$$

或对

$$x^4+y^4=z^4, \tag{7 b}$$

或更一般的对

$$x^n+y^n=z^n \quad 当\ n>2。 \tag{7 c}$$

作类似的讨论。费马断言，方程(7 c)在 $n>2$ 时，没有正整数解。这个命题既没有被证明，也没有被推翻，它叫作费马定理，或叫作费马大定理，以区别于另一个费马定理，该定理将在第 23 篇中提到。这个断言对于 n 的一定值已经得到证实。例如库马和他的后继者已经证明，由 3 到 100 的所有的 n 定理是对的。在此以前，欧拉已经对(7 a) 和 (7 b)作了证明。

利用我们关于毕达哥拉斯数的知识，可以很容易证明(7 b)没有正整数解。我们甚至能证明方程

$$x^4+y^4=w^2, \tag{8}$$

没有正整数解。因为每一个四次幂是平方数，但不是每个平方数是四次幂，因此(8)的不可解性较(7 b)更大。

在讨论解的过程中，我们坚持正整数的目的是为了消去一定的平凡解。例如 $x=1$，$y=0$，$w=1$ 一定满足(8)，而 $x=-y$，$z=0$ 满足(7 a)

8．如果我们把(8)写成如下形式

$$(x^2)^2+(y^2)^2=w^2, \tag{9}$$

可见它是(1)在 $a=x^2$，$b=y^2$，$c=w$ 时的特殊情形。我们只需考虑"简化解"x, y, z。可以看到(如上所述)，如果 (9) 有一个解，那么 w 必是奇数，而平方 x^2 和 y^2 必有一个是偶数，另一个是奇数。我们可以令 $x^2=a$ 是奇的，$y^2=b$ 是偶的。如果(9)有解，必可由(5)及(6)给出，这就必存在二个互素的，且是互反数对的素数 u 和 v，解由数

$$x^2=u^2-v^2, \tag{10 a}$$

$$y^2=2uv, \tag{10 b}$$

$$w=u^2+v^2, \tag{10 c}$$

确定。

现在(10 a)可以写为

$$x^2+v^2=u^2, \tag{11}$$

这是一个 x, u, v 为互素的新的毕达哥拉斯方程。其中 u 是在 c 的位置，所以它是奇的。因为 x 也是奇的，v 就必是偶的，并且是在 b 的位置。这个毕达哥拉斯方程(11)的简化解 x, v, u 将借助新的互素的互反数对的数 u_1 和 v_1，并通过(5)和(6)给出，即 x, u, v 由

$$x=u_1^2-v_1^2, \ v=2u_1v_1, \ u=u_1^2+v_1^2。 \tag{12}$$

给出。

我们现在回到方程(10 b)。因为 u 是奇数，v 是偶数，并且二数是互素的，并且由于 2 不能整除 u，可见 u 和 $2v$ 互素。因此(10 b)表示 y^2 是二个互素的因子 u 和 $2v$ 的乘积。根据第四节的讨论，一个平方数的因子分解，只有在它的互素因子都是平方数时才有可能。因而我们有

$$u=w_1^2, \ 2v=4t_1^2, \tag{13}$$

其中我们已利用了 $2v$ 是偶数的事实。如果我们把这二个值代入(12)中的后二个方程中，得

$$t_1^2=u_1v_1, \ w_1^2=u_1^2+v_1^2。 \tag{14}$$

现在 u_1 和 v_1 是互素的，并且它们的乘积是 t_1^2，所以它们必仍然是平方数，

$$u_1=x_1^2, \ v_1=y_1^2, \tag{15}$$

(14)的第二个方程现在成为

$$x_1^4+y_1^4=w_1^2。 \tag{16}$$

9. 方程(16)和原方程相同。由(8)的一个简化解 x, y, z (正整数)开始，我们求得同一方程的另一个简化解。这第二

个解也是第一个解通过一定的过程获得的。不完成整个过程，我们也能看到第一个解的 w 值大于第二个解的 w_1 值。因为，由(10 c)和(13)的第一个方程，有

$$w=u^2+v^2=w_1^4+v^2>w_1^4,$$

因此，$w>w_1$。这将使我们得到一个矛盾。正如刚才在第 8 节中我们可以由 x, y, w 得到 x_1，y_1，w_1 那样，我们现在也可以由 x_1，y_1，w_1 得到另一个解 x_2，y_2，w_2。如同 $w>w_1$ 一样，有 $w_1>w_2$。重复这个过程，将得到另一个解 x_3，y_3，w_3，且 $w_2>w_3$。连续不断的使用这个方法，我们得到

$$w>w_1>w_2>w_3>\cdots, \tag{17}$$

的一系列解。这些数都是正整数。只有有限个正整数小于 w，所以序列(17)必有最后的终点，设为 w_k。但是这个 w_k 是属于一个解 x_k, y_k, w_k 的。我们可以对它应用第 8 节中的程序来得到另一个解 x_{k+1}，y_{k+1}，w_{k+1}，且 $w_k>w_{k+1}$。这和(17)以 w_k 为终点的事实矛盾。因此我们确定了这样一个事实，即(8)没有任何正整数解，因为有这样一个解将得出一个矛盾来。

这个证明的基本思想是所谓的费马"无限减少原理"。它由寻求过程(在这里是反复应用第 8 节中的方法)所得到的矛盾组成。一方面，这个过程得到了一个没有终点的逐渐减少的正整数序列。另一方面，因为小于 n 的正整数只能是有限个，就是$(n-1)$，$(n-2)$，$\cdots$，3,2,1,即最多有$(n-1)$个，所以这样一个序列又必须有终点。

我们以前在第 4 篇中已经利用过一次无限减少原理，它在 $\sqrt{2}$ 的无理性的证明中是十分重要的。

[注] 即我国有名的勾股定理，最早见于"周髀算经"，历史上比毕达哥拉斯早五、六百年。毕达哥拉斯数在我国也称商高数。——译注

15. 算术平均和几何平均定理

一个细心的实验人员测量某个物体时，得到它的长度是2.172英尺。当他再作两次测量后，得到长度是2.176英尺和2.171英尺。他应当把哪个当作物体的真实长度呢？遇到这种情形，习惯上是采用测量值的平均值，也就是把测量值加起来，然后除以它们的次数。这个实验人员求出全部测量值应为6.519，然后除以3，求得2.173英尺，这就被当作是物体的长度。我们所说的这个平均值，叫作算术平均值。n个数a_1，a_2，a_3，……的算术平均值是

$$A=\frac{1}{n}(a_1+a_2+a_3+\cdots+a_n)。\tag{1}$$

我们将看到，除了算术平均值外，还有另一种类型的平均值，但是让我们首先看看平均值具有什么性质。如果实验人员把测量值都用英寸代替英尺，我们得到的平均值和以前得到的是一样的，只不过现在是用英寸来表示罢了。每个测量值都应当乘以12[1]，并且平均值也将乘以12。因为实验人员在度量时可以采用任意形式的单位，所以无论乘以12或是乘以任何其它数t，上述的情形都是对的。如果算术平均(1)是平均值的合理形式，它必具有如下性质：

如果每个a乘以t，那么最后的结果A也将乘以t。

$$tA=\frac{1}{n}(ta_1+ta_2+ta_3+\cdots+ta_n)。\qquad (2)$$

这显然是正确的，因为(2)的两端同除以 t，就得到等式(1)。我们用算术平均是“齐次的”来表示这个性质(2)。我们采用的任何平均值，都应当是齐次的。

平均值的第二个简单性质是这样的：a_1，$a_2,a_3\cdots\cdots a_n$ 的平均值既不小于最小的 a，也不大于最大的 a。实验人员若把 2.177 当作平均值是不合理的。算术平均值(1)之所以有这样的性质，是因为 n 个数的和，一定不小于这些数中最小值的 n 倍，并且也不大于这些数中最大值的 n 倍。

有一个浅显又简单的方法改变了算术平均值的形式。如果实验人员在作的三次测量中，最后一次较其它二次仔细，他感到应给第三个测量值 2.171 以更重的分量。如果他感到比其它值重要四倍，他应当列出头二个测量值 2.172 和 2.176 以及把第三值写上四次。现在平均值应当加上所有这些值，然后除以 6，求得 2.172 英尺。比较方便的是在列出最后一个值时记住要乘以 4，然后用 $1+1+4=6$ 除总数。这个平均值叫作加权算术平均值。如果已知 a_1 的权为 w_1，a_2 的权为 w_2，a_3 的权为 $w_3\cdots$，a_n 的权为 w_n，那么加权算术平均值是

$$W=\frac{1}{w_1+w_2+w_3+\cdots+w_n}(w_1a_1+w_2a_2+w_3a_3+\cdots+w_na_n)。$$

很容易验证，加权算术平均值具有一切平均值都有的两个性质。虽然我们将要讨论的内容可以推广到加权平均值，但我们将不作这个推广。我们只是引出了较为麻烦的算式，而没

有增加任何实质性的概念。

再举一例。假设我们有5个正方形，边长为1英尺的有二个，边长为2英尺的有一个，5英尺的有一个，7英尺的有一个。什么是正方形的平均值呢？如果我们感兴趣的是正方形边长的平均值，那么就是求长度的平均值，为3.2英尺。如果我们感兴趣的是正方形的面积，那么我们有二个正方形的面积是1平方英尺，一个是4平方英尺，一个是25平方英尺，一个是49平方英尺，面积的平均值就是16平方英尺。它对应于边长为4英尺的正方形。对于边来说，我们需要的平均正方形的边长为3.2英尺，但对于面积来说，平均正方形就要大些，其边长为4英尺。如果考察我们是如何得到这个数4的，那么看到首先是把长度平方，然后把这些平方数相加，最后除以它们的个数5，并就这个结果再取平方根。这样一个平均值叫作均方根。数 $a_1, a_2, a_3, \cdots a_n$ 的均方根是

$$R=\sqrt{\frac{1}{n}(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)}。\qquad (3)$$

这个平均值也具有一切平均值的两个性质。我们将只验证它是齐次的。如果我们用 t 乘(3)中的每个 a，有

$$\begin{aligned}&\sqrt{\frac{1}{n}(t^2a_1^2+t^2a_2^2+t^2a_3^2+\cdots+t^2a_n^2)}\\&=\sqrt{\frac{1}{n}t^2(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)}\\&=t\sqrt{\frac{1}{n}(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)}。\end{aligned}$$

这恰好是(3)的 R 乘以 t。

在前面这些例子中，算术平均值不大于均方根。这不是偶然的，它对任何数集 $a_1, a_2, a_3 \cdots a_n$ 都是能成立的。为了证明这个事实，首先要证明另一个事实(7)。证明这另一个事实的本身就是相当有趣的。

如果有一个数，无论是正，是负，或是零，它平方的结果或是正数或是零。因此，无论 c_1 和 d_1 是什么数，都有

$$0 \leqslant (c_1 - d_1)^2 = c_1^2 - 2c_1d_1 + d_1^2。$$

这个不等式两端同时加上 $2c_1d_1$，得

$$2c_1d_1 \leqslant c_1^2 + d_1^2。$$

类似的方法，有

$$2c_2d_2 \leqslant c_2^2 + d_2^2,$$

$$2c_3d_3 \leqslant c_3^2 + d_3^2,$$

$$\cdots\cdots$$

$$2c_nd_n \leqslant c_n^2 + d_n^2。$$

因为每个不等式左端都小于或等于右端，所以左端的和小于或等于右端的和，

$$2c_1d_1 + 2c_2d_2 + 2c_3d_3 + \cdots + 2c_nd_n \leqslant$$
$$c_1^2 + d_1^2 + c_2^2 + d_2^2 + c_3^2 + d_3^2 + \cdots + c_n^2 + d_n^2。$$

我们可以把它改写为如下形式

$$2(c_1d_1 + c_2d_2 + c_3d_3 + \cdots + c_nd_n)$$
$$\leqslant (c_1^2 + c_2^2 + c_3^2 + \cdots + c_n^2) + (d_1^2 + d_2^2 + d_3^2 + \cdots + d_n^2)。\quad (4)$$

这对于任意的数 $c_1, c_2, c_3, \cdots c_n, d_1, d_2, d_3 \cdots, d_n$ 都是能成立的。现在如果 $a_1, a_2, a_3, \cdots, a_n$ 和 $b_1, b_2, b_3, \cdots, b_n$ 也是任意数，我们可以取

$$r = \sqrt[4]{\frac{b_1^2 + b_2^2 + b_3^2 + \cdots + b_n^2}{a_1^2 + a_2^2 + a_3^2 + \cdots + a_n^2}}, \quad (5)$$

$$c_1=ra_1,\ c_2=ra_2,\ c_3=ra_3,\ \cdots,\ c_n=ra_n,$$

$$d_1=\frac{b_1}{r},\ \ d_2=\frac{b_2}{r},\ \ d_3=\frac{b_3}{r},\cdots,\ d_n=\frac{b_n}{r}。$$

因为$c_1d_1=r\cdot a_1\cdot\frac{b_1}{r}=a_1b_1,\ c_2\cdot d_2=a_2b_2,\ \cdots,\ c_nd_n=a_nb_n$，(4)式就成为

$$\begin{aligned}&2(a_1b_1+a_2b_2+a_3b_3+\cdots+a_nb_n)\\&\leqslant(r^2a_1^2+r^2a_2^2+r^2a_3^2+\cdots+r^2a_n^2)+\\&\quad\left(\frac{b_1^2}{r^2}+\frac{b_2^2}{r^2}+\frac{b_3^2}{r^2}+\cdots+\frac{b_n^2}{r^2}\right)。\end{aligned}\tag{6}$$

利用(5)式，右端的第一部分可以改写为

$$\begin{aligned}&r^2(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)\\&=\sqrt{\frac{b_1^2+b_2^2+b_3^2+\cdots+b_n^2}{a_1^2+a_2^2+a_3^2+\cdots+a_n^2}}(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)\\&=\sqrt{(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)(b_1^2+b_2^2+b_3^2+\cdots+b_n^2)}。\end{aligned}$$

类似地，(6)式的右端第二部分是

$$\begin{aligned}&\frac{1}{r^2}(b_1^2+b_2^2+b_3^2+\cdots+b_n^2)\\&=\sqrt{\frac{a_1^2+a_2^2+a_3^2+\cdots+a_n^2}{b_1^2+b_2^2+b_3^2+\cdots+b_n^2}}(b_1^2+b_2^2+b_3^2+\cdots+b_n^2)\\&=\sqrt{(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)(b_1^2+b_2^2+b_3^2+\cdots+b_n^2)}\end{aligned}$$

那么(6)式就成为

$$\begin{aligned}&2(a_1b_1+a_2b_2+a_3b_3+\cdots+a_nb_n)\\&\leqslant 2\sqrt{(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)(b_1^2+b_2^2+b_3^2+\cdots+b_n^2)}\end{aligned}$$

两端都除以 2，得

$$\begin{aligned}&a_1b_1+a_2b_2+a_3b_3+\cdots+a_nb_n\leqslant\\&\sqrt{(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)(b_1^2+b_2^2+b_3^2+\cdots+b_n^2)}。\end{aligned}\tag{7}$$

这就是以法国数学家柯西的名字命名的柯西不等式。

为了证明算术平均值不大于均方根，我们记住(7)式中的 a 和 b 可以是任意数，并且令 $b_1=b_2=b_3=\cdots=b_n=\frac{1}{n}$，那么我们有

$$\frac{a_1}{n}+\frac{a_2}{n}+\frac{a_3}{n}+\cdots+\frac{a_n}{n}\leqslant$$

$$\sqrt{(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)\left(\frac{1}{n^2}+\frac{1}{n^2}+\frac{1}{n^2}+\cdots+\frac{1}{n^2}\right)}。$$

其中右端后一个括弧中恰好有 n 项 $\frac{1}{n^2}$。由于 $n\cdot\frac{1}{n^2}=\frac{1}{n}$，不等式变为

$$\frac{1}{n}(a_1+a_2+a_3+\cdots+a_n)\leqslant\sqrt{\frac{1}{n}(a_1^2+a_2^2+a_3^2+\cdots+a_n^2)}。$$

因为左端是算术平均值(1)，右端是均方根(3)，所以这正是我们所需要的证明。

再回顾一下(7)式的证明方法是值得的。实质性的概念是数的平方永远不是负的。由此得到若干个不等式，它们的和为(4)式。虽然在(4)式中还看不到什么，但是，我们如果用另一些字母的一定的结合来代替 c 和 d 就得到了(7)。一组字母用另一些字母来代替的方法，在研究不等式时是经常采用的，有时甚至会带来意想不到的结果。(7)式的证明，虽然不是那么简单，但这是可以作出的，若首先证明了(7)，然后代换(7)的字母就得到了(4)。

为了介绍平均值的更多的形式，我们来看另一个衡量的例子。把一个物体放在天平的一个盘上，而在另一个盘上放一砝码使天平平衡，称得物体的重量。如果天平制造得不精

确，或者因为损伤而使天平二臂的长度略有不同，因而使称量不准确。要精确测定二臂长度通常是不可能的，所以我们将作二次称量，第一次是把要称量的物体放在左盘上，第二次再把它放在右盘上。如果这二次称量的结果是a_1和a_2，那么什么是物体的真正重量呢？我们是取算术平均值，还是另一种平均值更好呢？要回答这个问题，我们可以假设天平的左臂长度是l，而右臂长度是r。虽然我们不能准确测定这些长度，但是，臂必然有一定的长度。初等物理已经证明，重量和一端臂长的乘积，等于另一端的对应乘积。如果w是物体的真正重量，那么对于第一次称量就应有$wl=a_1r$，而对于第二次称量则是$wr=a_2l$。如果这些方程的左端相乘，同时右端也相乘，求得$wlwr=a_1ra_2l$，两端都除以lr，得$w^2=a_1a_2$。物体的真正重量就由公式$w=\sqrt{a_1a_2}$给出。这就是平均值的一个新形式。它就叫作数a_1和a_2的几何平均值。

多于两个数的几何平均值是指什么呢？如果我们已知n个正数$a_1, a_2, a_3, \cdots, a_n$，首先把它们乘在一起，得到$a_1a_2a_3\cdots a_n$。然后我们应求得这个数的平方根吗？要解决这个问题，我们应记得一个平均值必须是线性的。如果我们用t乘每一个a，乘积变为$ta_1\cdot ta_2\cdot ta_3\cdots ta_n=t^n\cdot a_1\cdot a_2\cdot a_3\cdots a_n$。如果将最后的结果乘以$t$，我们必须取$n$次方根，而不是平方根。我们现在可以说$n$个正数的几何平均值是

$$G=\sqrt[n]{a_1a_2a_3\cdots a_n}。 \qquad (8)$$

有关平均值的最著名的定理之一是几何平均和算术平均的定理。它表述为n个正数的几何平均值，不大于算术平均值，即$G\leqslant A$。这个定理有几种证明方法。我们将介绍的是由柯西给出的特别有趣而简单的证明。首先我们证明定理

为二个数时的情形。我们有

$$\left(\frac{a_1+a_2}{2}\right)^2=\frac{1}{4}(a_1^2+2a_1a_2+a_2^2)=\frac{1}{4}(a_1^2-2a_1a_2+a_2^2+4a_1a_2)$$

$$=\frac{1}{4}(a_1^2-2a_1a_2+a_2^2)+a_1a_2$$

$$=\left(\frac{a_1-a_2}{2}\right)^2+a_1a_2。$$

因为平方$\left(\frac{a_1-a_2}{2}\right)^2$不是负的，所以$\left(\frac{a_1+a_2}{2}\right)^2$是 a_1 a_2 增加一个非负的量。也就是

$$a_1a_2\leqslant\left(\frac{a_1+a_2}{2}\right)^2。\tag{9}$$

左端的平方根是G，而右端平方根是A，这就证明了任意二个数 a_1 和 a_2 情形时的定理。

现在我们证明个数更多时的定理。比较自然的是先证三个数，但我们暂时先来证明四个数时的情形。由于(9)对于任何数 a_1 和 a_2 都是正确的，我们可以用 a_3 代替 a_1，用 a_4 代替 a_2，有

$$a_3\cdot a_4\leqslant\left(\frac{a_3+a_4}{2}\right)^2。$$

由此以及(9)，我们有

$$a_1a_2a_3a_4\leqslant\left(\frac{a_1+a_2}{2}\right)^2\left(\frac{a_3+a_4}{2}\right)^2,$$

也即

$$a_1a_2a_3a_4\leqslant\left\{\left(\frac{a_1+a_2}{2}\right)\left(\frac{a_3+a_4}{2}\right)\right\}^2。\tag{10}$$

再一次利用对任何数都成立的(9)，但现在我们用$\left(\frac{a_1+a_2}{2}\right)$代替$a_1$，用$\left(\frac{a_3+a_4}{2}\right)$代替$a_2$，得

$$\left(\frac{a_1+a_2}{2}\right)\left(\frac{a_3+a_4}{2}\right)\leqslant\left\{\frac{\left(\frac{a_1+a_2}{2}\right)+\left(\frac{a_3+a_4}{2}\right)}{2}\right\}^2=$$

$$\left(\frac{a_1+a_2+a_3+a_4}{4}\right)^2。$$

把这用到(10)中，我们求得

$$a_1a_2a_3a_4\leqslant\left(\frac{a_1+a_2+a_3+a_4}{4}\right)^4。\tag{11}$$

因为左端的四次方根是G，而右端的四次方根是A，所以这就是四个数时我们的定理。

我们可以重复这些论证进而证明八个数时的情形。利用(11)，我们有

$$a_1a_2a_3a_4a_5a_6a_7a_8\leqslant\left(\frac{a_1+a_2+a_3+a_4}{4}\right)^4$$

$$\left(\frac{a_5+a_6+a_7+a_8}{4}\right)^4。$$

而且如同以前一样，可以把(9)应用于右端，得

$$a_1a_2a_3a_4a_5a_6a_7a_8\leqslant\left(\frac{a_1+a_2+a_3+a_4+a_5+a_6+a_7+a_8}{8}\right)^8,$$

这就给出了八个数时的定理。

如果我们用完全相同的方法继续证明下去，我们看到，如果$n=2,4,8,16\cdots=2,2^2,2^3,2^4,\cdots$亦即如果$n$是2的幂时，那么对于这样的$n$个正数$a_1,a_2,a_3,\cdots$的集合，有$G\leqslant A$。我们还必须证明$n=3,5,6\cdots$即不是2的幂时的定理。

现在，这是很容易的事。如果 $a_1, a_2, a_3, \cdots a_n$ 是 n 个任意正数，我们可以找到一个大于 n 的 2 的幂，设为 2^m（例如，如果 $n=50$，因为 $2^6=64>50$，我们取 $m=6$）。然后我们令

$b_1=a_1, b_2=a_2, b_3=a_3, \cdots, b_n=a_n$，$b_{n+1}=b_{n+2}=\cdots=b_{2^m}=A$，

其中 A 是 a 的算术平均值(1)。现在 $b_1, b_2, b_3, \cdots$，b_{2^m} 是2^m个正数的集合；我们早已证明，它们的几何平均值不大于它们的算术平均值，

$$\sqrt[2^m]{b_1b_2b_3\cdots b_{2^m}}\leqslant\frac{1}{2^m}(b_1+b_2+b_3+\cdots+b_{2^m})。$$

不等式的两端同时为 2^m 次方，并且把 b 的值代入，我们有

$$a_1a_2a_3\cdots a_nAA\cdots A\leqslant\left\{\frac{1}{2^m}(a_1+a_2+a_3+\cdots+a_n+A+A+\cdots+A)\right\}^{2^m}$$

由(1)，右端可化简为

$$\left\{\frac{1}{2^m}\cdot nA+\frac{2^m-n}{2^m}A\right\}^{2^m}=\left\{\frac{n+2^m-n}{2^m}A\right\}^{2^m}=A^{2^m}，$$

然后两端同除以 A^{2^m-n}，得到

$$a_1a_2a_3\cdots a_n\leqslant A^n。$$

两端再开 n 次方，有

$$\sqrt[n]{a_1a_2a_3\cdots a_n}\leqslant A=\frac{1}{n}(a_1+a_2+a_3+\cdots+a_n),$$

或 $G\leqslant A$，这就是算术平均和几何平均定理。

［注］ 1英尺＝12英寸　　——译注

16. 有限点集的跨度圆

1. 我们现在研究平面上由 n 个点 P_1, $P_2\cdots P_n$ 组成的有限集。每一对点 P_i 和 P_j 的距离都是能测量的。在这些距离组成的有限集中，必有最大者[1]，这最大距离称为点集的“跨度”。

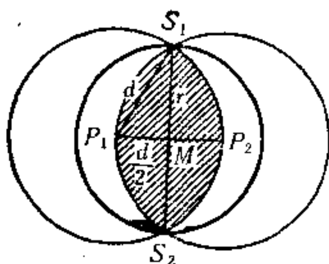


图 50

如果 n 个点的集的跨度为 d，那么我们能作一半径为 d 的圆，它完全包围了这 n 个点（图 50）。以这 n 个点的任一点比如 P_1 为中心，半径为 d 作圆，因为 P_1 到其它每一点的距离至多是 d，这个圆包含所有其它的 $P_2,P_3,\cdots P_n$。

然而，我们还能作一个仍然包围 n 个点的较小的圆。首先我们找一对点，它们的距离为 d，若这样的点对有若干，则任选一对，设为 P_1 和 P_2，然后以 P_1,P_2 为中心作二个圆。以 P_1 为中心的圆必过 P_2，反之亦然。现在该二圆中的任一个必包围了这个集的所有的点。这样，所有的点也就在平面上二圆的公共部分内，这部分就是图形中画阴影的地方。如果这二个圆交于 S_1 和 S_2，则以 S_1,S_2 为直径的圆包围了这公共面积，因而包围了整个点集。这新圆的半径 r，可通过把毕达哥拉斯定理用于三角形 P_1MS_1 而求得

$$r^2=d^2-\left(\frac{d}{2}\right)^2=\frac{3}{4}d^2,\ r=\frac{d}{2}\sqrt{3}。$$

我们称包围所有点的圆为该集的一个包围圆。原包围圆的半径为 $r=d$，我们刚建立的新圆的半径为 $r=\frac{d}{2}\sqrt{3}=0.866\cdots d$。

2．这个 $\frac{\sqrt{3}}{2}$ 能否用更小的数代替呢？荣格发现的下述定理给出了回答：任意一个跨度为 d 的有限点集，有一个半径不大于 $\frac{d}{3}\sqrt{3}=0.577\cdots d$ 的包围圆。有许多有限点集有更小半径的包围圆，而有些点集只有这样大的包围圆。另一方面，不论哪个包围圆内的任意二点相距不能比直径 $2r$ 更远，我们还有 $2r\geqslant d$，因而包围圆半径不能小于 $\frac{d}{2}$。荣格定理的证明就是本篇的目的。

对于边长为 d 的等边三角形的三个顶点组成的集，很容易求得半径为 $\frac{d}{3}\sqrt{3}$ 的一个包围圆，它是这个三角形的外接圆。利用图 51 的记号，且设 $r+x=h$，我们从直角三角形 ABD 中看出，有 $d^2=h^2+\frac{d^2}{4}$，所以

$$h^2=\frac{3d^2}{4},\qquad (1)$$

即

$$h=\frac{d}{2}\sqrt{3}。\qquad (2)$$

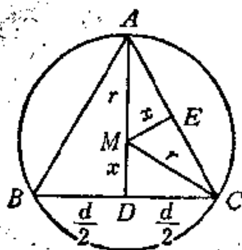


图 51

在三角形DCM中，我们有

$$x^2+\frac{d^2}{4}=r^2,$$

且有

$$(h-r)^2+\frac{d^2}{4}=r^2,$$

$$h^2-2hr+r^2+\frac{d^2}{4}=r^2,$$

$$h^2+\frac{d^2}{4}=2hr;$$

利用（1），我们得到

$$d^2=2hr,$$

$$r=\frac{d^2}{2h},$$

然后由（2）得 $r=\frac{d^2}{d\sqrt{3}}=\frac{d}{3}\sqrt{3}$。

对于等边三角形，这外接圆显然是最小包围圆。然而我们现在不详述这一点，因为今后它还要出现。

3．为了证明任意有限点集的荣格定理，我们将由试在所有可能的包围圆中选取最小半径的圆开始。我们将通过一系列步骤逐步得到较小的包围圆。

Ⅰ．一个包围圆C_1如果在其圆周上没有有限集S上的点，则通常能用一个较小圆C_2代替。我们可以以C_1的中心M为圆心，并通过离M最远的S的点(或某些点)作圆C_2。

Ⅱ．如果只有有限集S上的一个点在包围圆C_3上，那么这个圆可用一个较小的圆代替(图 52)。设P_1是在圆周C_3上S的一个点，我们作出通过S上的另一点且与C_3在P_1点有相同切线的所有圆。这些圆都在C_3的内部。设这些圆中

最大的一个为C_4。它与C_3不同，因为P_1是在C_3上的S的一点，而C_4有S的另一点在它的圆周上。现在C_4包含所有新圆及S的所有点，而且它有二个点在其圆周上且较C_3小。

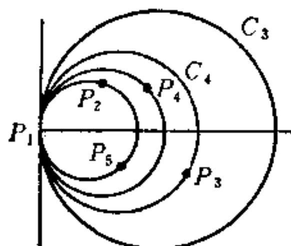


图 52

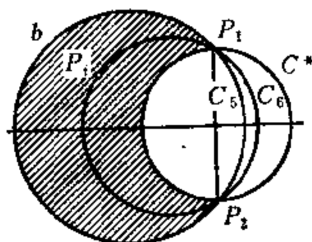


图 53

Ⅲ．在一包围圆周上的S的点能分割圆周为弧。为了方便，今后我们称除端点是S的点外，没有S的点在其上的弧为"自由——点"弧。我们缩小圆的第三个步骤可以这样来叙述：如果一包围圆的"自由——点"弧大过半圆周，那么该包围圆可由较小圆代替[8]。

令P_1和P_2是S的二个点，且是包围圆C_5的"自由——点"弧b的端点，且设b大过半圆周（图 53），我们以P_1P_2为直径作圆C^*，如果C^*包围了S的所有点，则它是较C_5小的包围圆，因为若C_5的弦P_1P_2不是直径（否则b是半圆周）那么C_5必大过C^*。如果不是所有的S的点都被C^*包围，那么剩下的点必是在b与C^*之间的新月形的面积内(图形的阴影面)。在b上没有不同于P_1和P_2的S的点，因为它是"自由——点"弧。我们作过P_1,P_2及新月形内的S的一点的所有的圆。每一个这样的圆的在C_5内部的那部分圆周是在新月形的内部，因而是在C^*的外部。在C^*内部的

那部分圆周是在 C_5 外部。设 C_6 是在新月形内的圆弧离弦 P_1P_2 最远的那个圆。那么这个圆 C_6 包围了 S 的所有的点，因为它包围了 S 在新月形内的所有的点，以及 C_5 和 C^* 的公共面积。这个面积包含了 S 的所有其它点。而且，C_6 较 C_5 小。它的圆周是在 C_5 和 C^* 的圆周之间，因而它的圆心较 C_5 的圆心距弦 P_1P_2 近，因此比较小。

如果一包围圆不再能利用Ⅰ.Ⅱ.Ⅲ而变小，那么它就不能有一个大过半圆周的“自由——点”弧，这样一个圆必然或者有 S 的二个点，它们是一个直径的端点，或者它必然有三个或更多个 S 的点，这些点分割圆周为若干段小于半圆周的弧。我们称第一类包围圆为“直径圆”，第二类为“三点圆”。把Ⅰ，Ⅱ，Ⅲ步应用于任一个包围圆，最终能得到这二类中的一种。包围圆也可能同时属于此二类，例如经过正方形四个端点的圆。

4．现在设想作出了所有以 S 的两个点为直径的端点的圆以及过 S 的三个点的圆[4]。不是所有这样的圆都是包围圆，但却包含了任意“直径圆”和“三点圆”。因为前面那些圆合起来是有限个，所以直径圆和三点圆只能有有限个。因而我们可以比较所有这些特殊的圆，并从中挑出最小的一个。这个圆 C 就是最小的包围圆。因为它是所有直径圆和三点圆中最小的一个，且由于Ⅰ，Ⅱ和Ⅲ，它必是小于任何其它包围圆的。并且它是唯一的。如果有同样大小的第二个包围圆 C'(图 54)，那么 S 将在 C 及 C' 内，于是 S 只能在它们的公共面积内，而这个面积可被较小圆 C^* 包围，这与 C 的极小性矛盾。我们将称这唯一确定的有限集 S 的最小包围圆为有限点集 S 的跨度圆。这跨度圆 C 上没有大过半圆周的“自

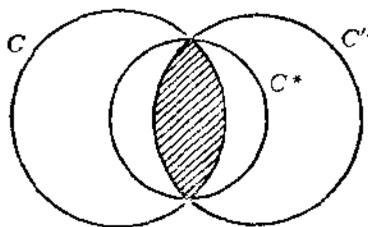


图 54

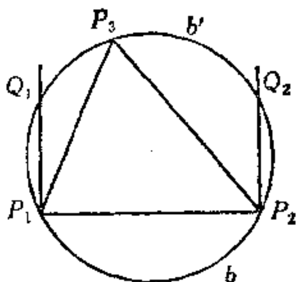


图 55

由——点弧”，因为根据Ⅲ，它就不是最小的包围圆了。

5．现在我们来证明跨度圆的半径不超过$\frac{d}{3}\sqrt{3}$。为此目的，我们于S在圆周C上的点中，选出相隔最远的点对。若这点对之间距离为δ，则一定不大于S的跨度d。

首先，可能碰巧这S的二个点是C的直径的端点。在此情况下，C的直径$2r$等于$\delta \leqslant d$。因而我们有$r \leqslant \frac{d}{2}$，从而肯定$r < \frac{d}{3}\sqrt{3}$。

其次，可能不是这种情况。那么我们选取圆周C上最大的“自由——点”弧b。如果有若干个弧都相等，则任选其一。b的终点P_1和P_2将是S的点。弧b小于半圆周，因为C上没有“自由——点”弧能大过半圆，而如果它恰等于半圆周时，则P_1和P_2就是直径的端点，我们也就将回到第一种情况。现在我们作弦P_1P_2，并在其上作直线垂直于它的端点（图

55)。这些垂直线将交圆于另外二点Q_1和Q_2。由Q_1和Q_2隔开的弧b'与弧b相对并和它全等。点Q_1和Q_2不属于S，因为Q_1和P_2正如Q_2和P_1一样，是直径的端点。因此，如果Q_1或Q_2属于S，我们将再次回到第一种情况。然而弧b'不可能是"自由——点"弧。因为，如果它是，它就能通过不属于S的Q_1和Q_2延伸，从而形成一个更大的自由——点弧。而这是不可能的，因为自由点弧不能大于b。

所以，在b'上的Q_1和Q_2之间，至少存在S的一个点P_3。点P_1，P_2，P_3组成一个锐角三角形。在P_1和P_2处是锐角，因为它们小于我们在这些点上所作的直角。在P_3上的角截取小于半圆的弧b。由于一个内接于半圆的角是直角，而且较小的角相当于较小的弧，所以P_3也是锐角。

圆周C被P_1，P_2，P_3分割为三段弧。其中一个弧必须至少和圆周的$\frac{1}{3}$一样大，但小于半圆周，因为P_1，P_2，P_3的角是锐角。因此它的弦P_iP_j必然至少和$\frac{1}{3}$圆周所对的弦一样大，即至少和内接于C的等边三角形的边s一样大。由于P_iP_j的长度至多是跨度d，我们有$s\leqslant d$。

我们在第二节已知外接于边为s的等边三角形的圆的半径为$r'=\frac{s}{3}\sqrt{3}$。由于$s\leqslant d$，关于C的半径r，我们有，$r\leqslant\frac{d}{3}\sqrt{3}$，这就完成了证明。

6. 我们还想知道，是否$\frac{d}{3}\sqrt{3}$的界限对一般的有限点集不能再减小了。让我们把第4节的方法用于由等边三角形

的三个顶点组成的点集T。如图 56 所示，作以T的二个点为直径端点的圆以及通过T的三个点所作的圆。这些圆中的唯一的一个包围圆是过三个点的圆，即等边三角形的外接圆。没有其他的圆能和这一个作比较，因此这就是跨度圆。这个三角形的边是跨度d，并且外接圆的半径是$r=\frac{d}{3}\sqrt{3}$。因为

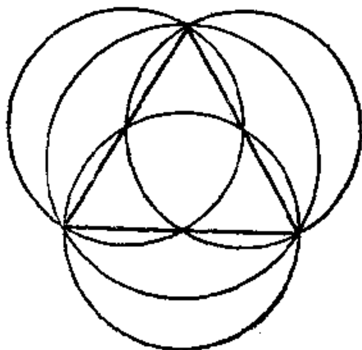


图 56

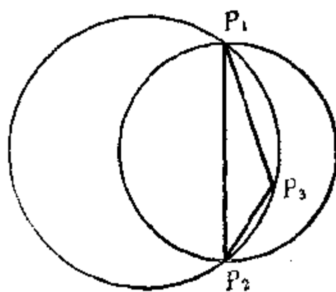


图 57

这个特殊的有限集有半径为$\frac{d}{3}\sqrt{3}$的跨度圆，所以一般有限点集的界限就是不能再减小的了。

这里，我们根据第4节证明了外接圆实际上是点集T的跨度圆。看起来这象是很明显的，但这并不是不证自明的。一个钝角三角形的顶点的跨度圆就不是外接圆，而是三角形的最大边作为直径的圆（图57）。

［注 1］利用第8篇第9节II，已知n个点中的不同的点对数为$\frac{n(n-1)}{2}$。

［注 2］I和II可看作是III的特殊情形，因为在这二种情形中整个圆周是"自由——点"弧。

［注 3］这些圆至多有$\frac{n(n-1)}{1\cdot2}+\frac{n(n-1)(n-2)}{1\cdot2\cdot3}$个。

17. 用有理数逼近无理数

π的一个历史悠久而熟知的近似值是$\frac{22}{7}$，（π是半径为1的圆的面积）。还有$\sqrt{2}$几乎是$\frac{7}{5}$。然而，这类说法的确切的数学含义是什么呢？"大约"和"几乎"这两个词在数学词汇中并无真正的地位，但这些说法必有某些意义。为什么$\frac{22}{7}$常常用来近似表示π呢？也就是说，为什么它总是比以8为分母的分数好呢？

1. 假定已知一个数w，那么就可以得出任意逼近它的

一些分数或者（如数学家常说的）一些有理数。例如，如果 $w=\pi=3.14159\cdots$，那么分数

$$3.1=\frac{31}{10},\ 3.14=\frac{314}{100},\ 3.141=\frac{3141}{1000},\cdots,$$

就越来越接近 π。很明显，第一个分数与 π 的差小于 $\frac{1}{10}$，因为 $\frac{32}{10}$ 已经是太大了，第二个分数与 π 的差小于 $\frac{1}{100}$，等等。用同样的方式，如果已知一个数 w 的小数展式，那么就能以任意精确程度来逼近数 w。

这个论断有一定的缺点，因为它与只是为了方便而随意选择的十进计算制联系着，而这种选择仅与数字的结构有关，并不反映数字的实质。然而，我们很容易从选择计算系统的偶然性中解脱出来，并且提出以下的论断，在这个论断中用任意一个整数 n 的幂代替 10，10^2 等等。

定理 1. **如果 w 是任意一个数，而 n 是任意整数，那么存在分母为 n 的有理数 $\frac{m}{n}$，它与 w 的差小于 $\frac{1}{n}$**，$0\leqslant w-\frac{m}{n}<\frac{1}{n}$.

例如，设 $w=\sqrt{2}$，$n=5$，那么 w 必在 1 与 2 之间，因此它处在数

$$1,\ \frac{6}{5},\ \frac{7}{5},\ \frac{8}{5},\ \frac{9}{5},\ 2, \tag{1}$$

所形成的 5 个区间中的某一个内。这些区间每一个的长都是 $\frac{1}{5}$。无需进一步计算就可知道，这些数中小于 $\sqrt{2}$ 的最后的

那个数，是分母为5且与$\sqrt{2}$的差小于$\frac{1}{5}$的分数。一般，对于任意给定的w和整数n，需要考察有理数

$$g,\ g+\frac{1}{n},\ g+\frac{2}{n},\cdots,g+\frac{n-1}{n},\ g+1, \tag{2}$$

这里g是小于w且最接近w的整数。在这些数中必有小于w的最后的数，或等于w的数，因而与w的差应小于$\frac{1}{n}$。设这个数等于$g+\frac{l}{n}$，那么有不等式

$$0\leqslant w-\left(g+\frac{l}{n}\right)<\frac{1}{n} \tag{3}$$

成立。这就是要证明的定理Ⅰ。

让我们来寻找$\sqrt{2}$介入其中的数列(1)中的两个数。为了计算的方便，设法消去分母5，于是将数列(1)的所有数增加5倍。问题转化为，在数列(1)的5倍数的哪个区间内，即在数列5，6，7，8，9，10之间的哪个区间内，恰好落入$\sqrt{2}$的五倍值，即$5\sqrt{2}=\sqrt{25}\sqrt{2}=\sqrt{50}$,或者换句话说，要找的是小于$\sqrt{50}$的最大整数。因为$49<50<64$，那么$7<\sqrt{50}<8$；这样，找到的数是7。除以5,可知$\sqrt{2}$在$\frac{7}{5}$和$\frac{8}{5}$之间，于是得

$$0\leqq\left(\sqrt{2}-\frac{7}{5}\right)<\frac{1}{5}。 \tag{4}$$

如果我们也消去分母，定理Ⅰ的证明也可以大大简化。

我们以 nw 代替 w 加以考察，并且寻求不超过 nw 的最大整数。如果 m 是所找到的不超过 nw 的最大整数，则有 $m \leqslant nw < m+1$,

因此 $$0 \leqslant nw-m < 1。$$

除以 n，得

$$0 \leqslant w-\frac{m}{n} < \frac{1}{n}.$$

这个不等式正是我们所需要的结果。

2. 因此，关于确有有理数接近 $\sqrt{2}$ 和 π 这一事实，并没有什么特别的地方。那么 $\frac{22}{7}$ 近乎等于 π，$\frac{7}{5}$ 近乎等于 $\sqrt{2}$ 的论断是什么意思呢？这就是表示 $\frac{7}{5}$ 大致地接近 $\sqrt{2}$，它们的差小于 $\frac{1}{5}$。这是刚才证明的定理所保证的。但是，可以有更精确的估计。验证下述不等式是容易的：

$$\frac{7}{5} < \sqrt{2} < \frac{17}{12}.$$

因此，$\sqrt{2}$ 与 $\frac{7}{5}$ 的差小于 $\frac{17}{12}$ 与 $\frac{7}{5}$ 的差，即

$$\sqrt{2}-\frac{7}{5} < \frac{17}{12}-\frac{7}{5}=\frac{1}{60}.$$

但由定理 1 得到的则是差应小于 $\frac{1}{5}$。与此类似，阿基米德给出了如下不等式

$$3+\frac{10}{71} < \pi < 3+\frac{1}{7},$$

从而断定，

$$\frac{22}{7}-\pi<\left(3+\frac{10}{70}\right)-\left(3+\frac{10}{71}\right)=\frac{10}{70}-\frac{10}{71}$$

$$=\frac{10}{70\cdot 71}=\frac{1}{497},$$

而按定理 I 得到的这个差的界等于$\frac{1}{7}$。

但是，有理数“更加接近”已知数的思想仍未成为一种数学概念。现在要证明一个给这个问题以完全确定意义的定理。

定理 2. 如果w是一个无理数，而N是任一个整数，那么就有一个分数$\frac{m}{n}$，它的分母不超过N，而它与w的差小于$\frac{1}{Nn}$并且还有无限多个与w的差小于$\frac{1}{n^2}$的分数$\frac{m}{n}$.

应用上面的二个例子，可以断定$\sqrt{2}-\frac{7}{5}$小于$\frac{1}{5^2}=\frac{1}{25}$，而$\frac{22}{7}-\pi$小于$\frac{1}{7^2}=\frac{1}{49}$。虽然实际的差仍小于这个新定理所确定的界。但这个定理较第一个极平常的定理大为精确，因而后者是前者的一个改进。

为证明这个定理，我们考察的不仅是Nw和小于它的最大整数，而是一系列数

$$w,\ 2w,\ 3w,\cdots,\ Nw,$$

以及小于上面这些数的最大整数：g_1, g_2, g_3, $\cdots$, g_N。于是我们有

$$0<w-g_1<1,\ 0<2w-g_2<1,\ \cdots,\ 0<Nw-g_N<1,$$

为了便于理解，我们先举一特例。设 $w=\sqrt{2}, N=13$，得

$$\begin{aligned}
\sqrt{2}&=1.414\cdots=1+0.414\cdots\\
2\sqrt{2}&=2.828\cdots=2+0.828\cdots\\
3\sqrt{2}&=4.242\cdots=4+0.242\cdots\\
4\sqrt{2}&=5.656\cdots=5+0.656\cdots\\
5\sqrt{2}&=7.071\cdots=7+0.071\cdots\\
6\sqrt{2}&=8.485\cdots=8+0.485\cdots\\
7\sqrt{2}&=9.899\cdots=9+0.899\cdots\\
8\sqrt{2}&=11.313\cdots=11+0.313\cdots\\
9\sqrt{2}&=12.727\cdots=12+0.727\cdots\\
10\sqrt{2}&=14.142\cdots=14+0.142\cdots\\
11\sqrt{2}&=15.556\cdots=15+0.556\cdots\\
12\sqrt{2}&=16.970\cdots=16+0.970\cdots\\
13\sqrt{2}&=18.384\cdots=18+0.384\cdots
\end{aligned}$$

在这 13 个数分别与相应的最接近的整数的差的余数中，第 5 个是最小的，$5\sqrt{2}-7=0.071\cdots$，第 12 个最大：$12\sqrt{2}=16+0.970\cdots=17-0.030$，或 $17-12\sqrt{2}=0.030$。假设这 13 个余数在我们的表中忽增忽减，我们按它们数值的大小排列起来；于是在 0 与 1 的间隔中，将有分割此间隔为 14 个区间的 13 个数。当然，这些间隔是不相等的，但至少有一个小于$\frac{1}{14}$。因为如果它们都$\geqslant\frac{1}{14}$，那么总和至少是$\frac{14}{14}=1$，并且只有当每一个都恰好是$\frac{1}{14}$时，它们的总长度才是 1。

但如果它们都恰好是$\frac{1}{14}$，那么$\sqrt{2}$就是1加上13个余数中的某一个所得的数，而这些余数都是$\frac{1}{14}$的倍数，因而$\sqrt{2}=1+\frac{m}{14}$。也就是说，$\sqrt{2}$将是有理数。由于我们假定w是无理数，并且已在第4篇中证明$\sqrt{2}$的无理性，从而得出结论，这些区间不可能都是$\frac{1}{14}$，其中至少必有一个小于$\frac{1}{14}$。我们只知道有这样的区间存在，但不知道是哪一个。设$a\sqrt{2}-g_a=r_a$是这个区间的下端点，而$b\sqrt{2}-g_b=r_b$是上端点，因而

$$0<r_b-r_a=(b\sqrt{2}-g_b)-(a\sqrt{2}-g_a)<\frac{1}{14},$$

由此，
$$0<(b-a)\sqrt{2}-(g_b-g_a)<\frac{1}{14}。$$

由于其中a和b是数列0，1，···，13中的某二个数，且a和b不等，则差$b-a$的绝对值是这些数中之一，即$-13\leqslant b-a\leqslant 13$。因此$(b-a)\sqrt{2}$或$(a-b)\sqrt{2}$中有一个是正的，是13个$\sqrt{2}$的倍数中的一个，我们把它称作$n\sqrt{2}$。且$n\sqrt{2}$与整数（$g_b-g_a$或$g_a-g_b$）的差的绝对值小于$\frac{1}{14}$。用$m$表示这个整数，则

$$-\frac{1}{14}<n\sqrt{2}-m<\frac{1}{14},\ n\leqslant 13,$$

除以n，得到最后的结果

$$-\frac{1}{14n}<\sqrt{2}-\frac{m}{n}<\frac{1}{14n}。$$

一般的证明可按完全相同的方法进行。如果用无理数w代替$\sqrt{2}$，N代 13，$N+1$ 代 14，那么上式就成为：存在这样的$n \leqslant N$，对于它，不等式

$$-\frac{1}{(N+1)n}<w-\frac{m}{n}<\frac{1}{(N+1)n} \qquad (5)$$

成立。这就是定理 2 的第 1 部分。

因为 $n \leqslant N$，那么由（5）立即得到

$$-\frac{1}{n^2}<w-\frac{m}{n}<\frac{1}{n^2}, \qquad (6)$$

这样我们找到了与 w 的差小于$\frac{1}{n^2}$的分数。在不等式(6)中，数N并未出现，但它对于求分数$\frac{m}{n}$是有作用的，因为分数的分母 n 需要 $n \leqslant N$。

在前面的例子中，当N是 5，6，···，11 中的任一个时，数 n 等于 5，即

$$0<5\sqrt{2}-7<\frac{1}{5}, \quad 0<\sqrt{2}-\frac{7}{5}<\frac{1}{5^2}。$$

当N达到 12 时，n 成为 12，于是我们有

$$0<17-12\sqrt{2}<\frac{1}{12}, \quad 0<\frac{17}{12}-\sqrt{2}<\frac{1}{12^2}。$$

如$N=13$，n 仍是 12 等等。

定理 2 的第二部分是，要求对于任一已知的无理数 w，证明存在无限多个满足不等式（6）的分数$\frac{m}{n}$。这只要证明对

于满足(6)的任意一个分数$\frac{m}{n}$，一定能找到另一个分数$\frac{m'}{n'}$，它比$\frac{m}{n}$更接近 w，且同样满足不等式(6)。由于差 $w-\frac{m}{n}$不能等于 0(w 是无理数，它不等于任意的分数)，那么它与 0 的差是一确定的数。于是存在分母是足够大的分数$\frac{1}{N'}$，它与 0 的差小于 $w-\frac{m}{n}$，即

$$0<\frac{1}{N'}<w-\frac{m}{n}, \text{ 或 } 0<\frac{1}{N'}<\frac{m}{n}-w. \tag{7}$$

究竟是上面哪一式，则视差 $w-\frac{m}{n}$是正或负而定。对于这个数 N'(就如同对于数N一样)，我们可以按刚才所证明的方法求得分数$\frac{m'}{n'}$，因而在 $n'\leqslant N'$的条件下，满足了类似于(5)的不等式

$$-\frac{1}{(N'+1)n'}<w-\frac{m'}{n'}<\frac{1}{(N'+1)n'}. \tag{8}$$

由此得到需要的不等式

$$-\frac{1}{n'^2}<w-\frac{m'}{n'}<\frac{1}{n'^2}。$$

而且，由于(8)是真的，我们当然就有

$$-\frac{1}{N'+1}<w-\frac{m'}{n'}<\frac{1}{N'+1},$$

即 $w'-\frac{m'}{n'}$与 0 的差**小于**$\frac{1}{N'+1}$。但根据(7)，$w-\frac{m}{n}$与 0 的差大于$\frac{1}{N'}$。因此$\frac{m'}{n'}$较$\frac{m}{n}$更接近 w；所以$\frac{m}{n}$与$\frac{m'}{n'}$是不同

的。

3．因此不存在满足不等式(6)的最后一个分数$\frac{m}{n}$。对于任意这样的分数都有随后的一个，这就是说，正如定理2所断言的，有无限多个。同时，在极平常的定理1中，允许任意整数作分母，而这个定理中，选取的是特殊的分母，仅是具有性质(6)的整数。我们可以称这样的分母为“最佳分母”。在前面一些例子中，当逼近$\sqrt{2}$时，2，5，12是最佳分母。进一步的计算得到最佳分母是29，70，169，…。

对于π来说，7是最佳分母。实际上还有比定理2所保证的更好的处理方法。因为$\frac{22}{7}-\pi<\frac{1}{497}$，大大小于$\frac{1}{7^2}=\frac{1}{49}$。这就产生一个问题，对于任何一个数，有没有比定理2更优的分母，例如满足$-\frac{1}{n^3}<w-\frac{m}{n}<\frac{1}{n^3}$，或某些类似的不等式。

我们现在要指出，定理2的改进，对任何无理数w一般是不可能的。我们将考虑特殊的值$w=\sqrt{2}$，并将指出，任意一个分数$\frac{m}{n}$，它与$\sqrt{2}$的差总是大过$\frac{1}{3n^2}$的。

我们首先考察大于$\sqrt{2}$的分数。当$\frac{m}{n}\geqslant 2$时，因为$\sqrt{2}<1.45$，则差$\frac{m}{n}-\sqrt{2}$大于0.55，而$\frac{1}{3n^2}\leqslant\frac{1}{3}$小于0.55。当$1.55\leqslant\frac{m}{n}<2$时，因为$\sqrt{2}<1.45$，所有$\frac{m}{n}\geqslant 1.55$的分数与$\sqrt{2}$的差大过0.1，而$n$至少是2，所以$\frac{1}{3n^2}$从$n=2$

开始就小于 0.1；这样，对于$\frac{m}{n}\geqslant 1.55$，我们的判断是正确的。现在如果$\frac{m}{n}$在$\sqrt{2}$和 1.55 之间，那么

$$\left(\frac{m}{n}\right)^2-2=\frac{m^2-2n^2}{n^2}=\frac{g}{n^2},$$

其中 g 是某个正整数，这就是说，g 至少是 1，而且

$$\left(\frac{m}{n}\right)^2-2\geqslant\frac{1}{n^2}。$$

由公式 $a^2-b^2=(a+b)(a-b)$得

$$\left(\frac{m}{n}+\sqrt{2}\right)\left(\frac{m}{n}-\sqrt{2}\right)\geqslant\frac{1}{n^2},$$

由此

$$\frac{m}{n}-\sqrt{2}\geqslant\frac{1}{n^2}\cdot\frac{1}{\frac{m}{n}+\sqrt{2}}。$$

由于$\frac{m}{n}<1.55$，我们得出$\frac{m}{n}+\sqrt{2}<1.55+1.45=3$，于是，这个和数的倒数大于$\frac{1}{3}$，因此

$$\frac{m}{n}-\sqrt{2}>\frac{1}{3n^2},$$

这正是我们的论断。

最后，如果 $0<\frac{m}{n}<\sqrt{2}$，我们同样可以推断：

$$2-\left(\frac{m}{n}\right)^2=\frac{g}{n^2}\geqslant\frac{1}{n^2},$$

$$\sqrt{2}-\frac{m}{n}\geqslant\frac{1}{n^2}\frac{1}{\frac{m}{n}+\sqrt{2}}>\frac{1}{n^2}\cdot\frac{1}{2\sqrt{2}}>\frac{1}{3n^2}。$$

18．利用连杆产生直线运动

1．瓦特原来的蒸汽机安装了一个值得注意的，按其形状而称为瓦特平行四边形的机械结构。这个结构草图如图58，由绞结在 C、D、E、F 的五个杆组成。在 A 和 B 处，杆与轴相连，杆因而能旋转。在这个草图中，用黑圈表示枢轴的固定位置，用空心圆圈表示可移动的枢轴。虚线是辅助线并不表示杆。活塞杆附在 F 上。这种结构用来使力作用在活塞终点，且作直线运动，从而使活塞杆能塞进汽缸。但从数学上看，F 点并不是按严格的直线而是按近乎直线的曲线运动的。这也就是所发明的机械结构的用处。

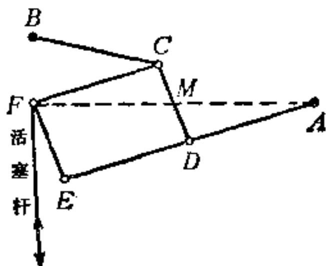


图 58

平行四边形 $CDEF$ 是取名为连杆的一部分，它不是结构的关键部分。它只是用来扩大运动的可利用部分。其关键部分是连杆 $ADCB$。对于不是很大的运动来说，这个部分使 DC（杆 AE 在 D 点不弯曲）的中点 M 近似沿着直线运动。如果连杆按照 $AD=DE=CF=BC$ 和 $DC=EF$ 制成（为了使汽缸中的活塞能伸缩这是必要的），那么由于$\triangle AEF$ 和$\triangle ADM$ 相似，点 A、M、F 将一定在一条直线上，因而 AF 较 AM 放大二倍，点 F 的运动“相似”于点M的运动。

我们也能把M点的运动放大K倍。我们保留$AD=BC$，这样CD的中点M会近似沿直线运动，但我们增加$\square DEFG$的尺寸，(图 59)得到$AD:AE=DM:FE=1:K$。

为了研究瓦特平行四边形实际产生的运动，我们现在只限于讨论M点的运动。只要考察"连杆"的几个极端位置，就能看出这不是直线运动。但是，由于这个具体运动的重要性不如我们另外要讲的题目，我们就不进一步去研究它。

2. 研究连杆产生直线运动的问题，在机械设计史中有

图 59

其重大意义，并且人们对这个问题作了大量的工作。但是，瓦特的连杆已不再受到重视了。在现代热力工程中，利用的是完全不同样式的机械结构，有一个"丁字头"在平行横轨之间滑动，使活塞杆的端点按直线运动。可能是对于摩擦力的不正确的观念，使设计者当时拒绝了丁字头而想方设法利用连杆。

连杆问题不仅受到实际设计人员的关注，同时也受到了数学家的注意。数学家自然考虑的是结构形式的问题，即找一个连杆，其中的一点是按严格的直线运动的。最先研究此问题的是伟大的俄国数学家契比雪夫，他研究了瓦特的平行

四边形，以及找一个能产生精确的直线运动的连杆的问题，但可能没有获得成功。在上个世纪的下半叶，进行了许多徒劳的努力，后来，数学家甚至开始怀疑严格的解是否可能。

然而，在1864年，皮塞勒设计了一个产生直线运动的连杆。这个装置称为皮塞勒网。象通常在发明史中所遇到的，随后找到了这个问题的大量的解。可以产生各种曲线运动的连杆是能找到的，而直线只不过是曲线中的特殊情形。

3．现在我们来研究皮塞勒网。这个连杆不仅产生直线运动，同时还有一个平面的"反演"。反演是一种平面到平面的特殊的"映射"。映射的意思是对于平面图形的每一个点都有一个像。一些最简单的映射是：(1) 直线的反射。这已在第5、6篇中反复应用过了。(2) 平行移动。就是每个点的像是原来的点按确定的方向平行移动一个常量而得到。(3) 平面关于定点旋转一个角度的映射。(4) 关于一个中心点的伸缩映射。通过这个映射，每个点沿中心发出的射线运动，到达中心的距离时将增加一个固定的比 $1:\lambda$。一个点的像不一定是与原点不同的点。在平移中，每个点映射后的像是不同的，但在旋转或伸缩中，中心的像就是中心本身。在直线的反射中，直线上的每个点都被映射到自身。

这些例子可以用来说明映射的意思是什么。现在我们还可看到另一个映射——反演。如图60所示，圆 C 以 O 为中心，这也是反演的中心。为了找出任一点 P 的像 P'，我们从 O 过 P 作一射线。这射线交圆于某点 Q。我们使在射线上的 P' 满足 $OP:OQ=OQ:OP'$。如果圆的半径为 a，并写成 $OP=r, OP'=r'$，那么，其比例可由方程 $r\cdot r'=a^2$ 代替。

按照反演的定义，显然在反演圆 C 内部的每一个点的像

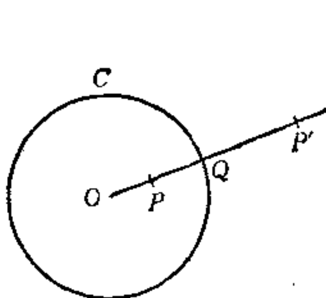


图 60

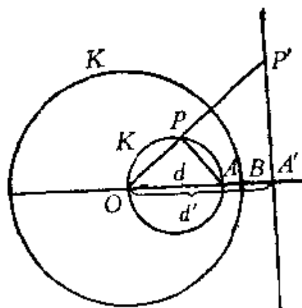


图 61

是在圆的外部，反之亦然。反演圆上的每一点都是它自己的像。由此，正如普通的反射一样反演通常也称为"圆的反射"。反演有这样的性质：像的像就是原来点的本身。

4．迄今为止，有关单独的点的所有反演的性质，都是从定义直接证明的。如果我们讨论曲线的像，那么就产生了新问题，而且也不能担保这个问题能完整解决。最重要的结果是，每一直线或圆经反演后仍是圆或直线。本篇我们将只用到这个结论的一部分，即**过反演中心的圆的像是一直线**。我们现在证明这个论断。

设K是过反演中心O的圆，并设它的直径是$OA=d$(图61)。反演圆K的半径为a。A的像A'在直径OA的延长线上，并且距离$OA'=d'$必须满足方程$dd'=a^2$。现在证明K的像是一条过A'与OA'垂直的直线。为此，必须证明过O的每一条射线交圆K于点P，交该直线于P'，且$OP\cdot OP'=a^2$。如果我们作直线AP，我们有二个直角三角形OAP和$OP'A'$，由于它们在点O有公共角，所以是相似的。因此我们有比例

$$OP:OA=OA':OP',$$

或用 $OP=r, OP'=r'$ 代入，有

$$r\cdot r'=d\cdot d'.$$

但利用 $d\cdot d'=a^2$，我们有

$$r\cdot r'=a^2.$$

这恰好是反演所要求的点和其像的关系。

和这个事实一起，我们还需要一个在初等几何中已证明的结果。如果过圆外一点作圆的二条割线，那么一条直线和它的圆外一段的乘积，等于另一条直线与它的圆外的一段的乘积。在图 62 的记号中，这个定理可以写为 $S_1S_1'=S_2S_2'$。它的证明可由三角形 AP_1P_2' 和 AP_2P_1' 相似直接得到。这二个三角形相似，是由于有公共角 A，且在 P_1' 和在 P_2' 处的角相等，因为它们对同弧 P_1P_2。

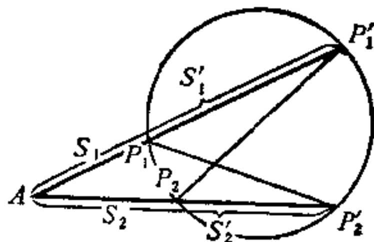


图 62

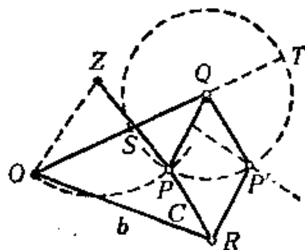


图 63

5．在作了这些准备后，现在回过来讨论皮塞勒网（图 63）。它由四根长为 C 的杆绞结为菱形 $PQP'R$，并由长为 $b>c$ 的二根杆绞结在菱形的二个对点 Q 和 R 而组成。二根杆的另二个端点都接在作为枢轴的定点 O 处。

由于这个杆的系统的对称性，点O，P，P'经常是在一条直线上，并且这还是图形的对称轴。以Q为心，c为半径作圆。这个圆通过P和P'，且交OQ及其延长线于S，T二点。那么OPP'与OST就是圆的二条割线，对于它们可应用上面提到的初等定理，给出结果$OP\cdot OP'=OS\cdot OT=(b-c)(b+c)=b^2-c^2$。因此，乘积$OP\cdot OP'$是常数。如果令$a^2=b^2-c^2$，那么$a$是斜边为$b$，另一边为$c$的直角三角形的一边，那么，由$OP\cdot OP'=a^2$知$P'$是$P$关于以$O$为心以$a$为半径的圆的反演。

因而皮塞勒网建立了一个平面部分的反演。为了得到直线运动，只要使P是在过O点的圆上运动即可。从反演的已知性质，则像P'就沿直线运动。为使P在过O的圆上运动，只要在P再绞结一个杆，使其终点在固定点Z上，并在此安装一轴。为了作过O的圆，我们必须令杆ZP的长度，等于固定的距离OZ。

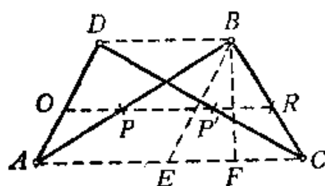


图 64

由于加了一个杆，皮塞勒网由七杆组成。P点将限制在一个过O的圆的圆弧上，则P'沿直线运动。这条线将垂直于OZ。

6. 已经找到另一些连杆也能产生反演，且能用来产生直线运动（如同皮塞勒网一样）。其中之一是由哈特设计的，只要求用四个杆，代替了皮塞勒网的六个杆。这四个杆绞结在一起，组成平行四边形$ABCD$，但却是“内外交错”的。即把点A和C拉开直到平行四边形塌下，AC成直线以前，

D和B都在AC的同一侧（图 64)。二个“对角线”AC和DB总是平行的，因为三角形ACB和ACD是全等的，所以高度相同。在这个连杆处于某个确定位置时，能在每个杆上分别找到点O, P, P', R，它们在平行于这些“对角线”的直线上。现在我们要证明，这些点对于连杆的一切其它位置，仍是在平行于“对角线”的直线上。

首先看图 64 的三角形DAB，我们注意到OP是平行于DB的，这DB是连杆的初始位置。由此，我们有

$$AO:OD=AP:PB. \tag{1}$$

我们所证明的(1)只是对于连杆的初始位置成立，但因为(1)中的比例只包含线段长度，所以对于连杆的任何位置必也是成立的。从(1)我们有$OP \parallel DB$，并且对于连杆的任何位置都成立。现在再看三角形ADC，我们发现OP'是平行于连杆的初始位置的，因此，我们有

$$DO:OA=DP':P'C, \tag{2}$$

这个线段长度之间的比例，对于连杆的任何位置都是对的，因此我们有$AC \parallel OP'$。由于我们已经证明$OP \parallel DB$, $DB \parallel AC$, $AC \parallel OP'$，因此我们有$OP \parallel OP'$，而且由于它们有公共点O并且平行，所以OP和OP'是同一条直线。这条线平行于DB。类似的论证可用于三角形DCB，从而证明$P'R \parallel DB$，并求得R也是在包含O, P, P'的直线上。

而且，由于三条直线是平行的，我们有

$$OP:DB=AO:AD,$$

$$OP':AC=DO:DA。$$

这些比例可由以下方程代替：

$$OP \cdot AD=AO \cdot DB, \tag{3}$$

$$OP'\cdot DA=DO\cdot AC。\tag{4}$$

如果我们把以上方程的对应边相乘并除以 AD^2，得到：

$$OP\cdot OP'=\frac{AO\cdot DO}{AD^2}\cdot AC\cdot DB。\tag{5}$$

右边的数 AO, DO 和 AD 是由连杆确定的固定长度，AC 和 DB 则是腰的长度，表面上依赖于连杆的位置，但 $AC\cdot DB$ 的积是常数。因为作 $BE\parallel DA$，和 $BF\perp AC$，我们有

$$AC\cdot BD=(AF+FC)(AF-FC)=AF^2-FC^2\tag{6}$$

由直角三角形 AFB 和 FCB 及毕达哥拉斯定理，我们求得

$$AF^2+FB^2=AB^2,$$

$$FC^2+FB^2=CB^2,$$

然后，它们相减 $AF^2-FC^2=AB^2-CB^2$。

这样，代入(6)，给出

$$AC\cdot BD=AB^2-CB^2。\tag{7}$$

由于右边是常数，所以不依赖于连杆的位置。现在由(7)和(5)，我们有

$$OP\cdot OP'=\frac{AO\cdot DO}{AD^2}(AB^2-CB^2)。\tag{8}$$

这里右边是常数，可由具体的连杆而定。由于 O, P, P' 在一直线上，则点 P' 是 P 对于固定位置 O 反演后的像，这个反演圆的圆心在 O 点，而其半径的平方由方程(8)的右端给出。

现在，哈特的连杆产生一个反演的证明已完成了。为了得到一个直线运动，只需要加上第五个杆，使 P 在过 O 的圆上运动即可。

7．也有不依赖于反演产生直线运动的连杆。肯普的二

重长菱形(图 65)就是这样一个特别巧妙的装置。所谓长菱形是指两组邻边分别相等的四边形。在长菱形（或筝形）中，分别由二个不等边组成的二个角是相等的。设 $ABCD$ 和 $BCEF$ 是二个铰结在一起的长菱形。这二个长菱形的小边与大边的比是相同的，而且大菱形的小边等于小菱形的大边。那么在图 65 中，我们有 $AD=AB, CD=CB=CE, FB=FE$ 和$CB:AD=FB:CB$。由于小菱形的顶点F在大菱形的一个边上，则二个菱形在B有公共角。因而在B，在E和在D的角都相等。现在这二个长菱形相似，这是因为它们的边有相同的比，且对应的一对角相等。并且无论连杆怎样运动，这两个菱形 $ABCD$ 和 $BCEF$ 总是相似的。

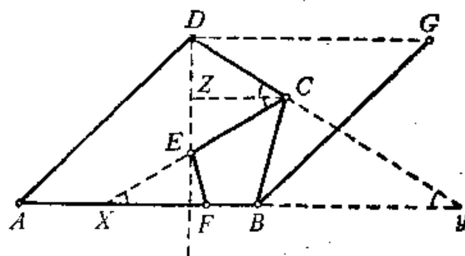


图 65

如果把长菱形的对边都延长，直到它们相交，它们就组成一个角。由于菱形的相似，这些角，EXF 和 CYB 显然相等。如果我们过C作 $CZ \parallel AB$,我们有$\angle DCZ=\angle CYB$，因为它们是平行线的同位角。我们还有 $\angle ZCE=\angle EXF$，因为它们是内错角。所以标出的四个角都是相等的。因而 CZ 是三角形 CDE 的顶角平分线，这 CDE 按假设是等腰

三角形。CZ 则垂直于底 DE。而 CZ 平行于 AB,所以DE 必然总是垂直于 AB。

现在如果我们令D固定，并且令 AB 按总是平行于初始位置的方式运动，那么由D垂直于 AB 的直线，也会是固定的。但点 E 是在这垂线上，所以就只能沿垂线作直线运动。为了使 AB 始终平行于它本身，我们增加了一个杆BG，其长度为 $BG=AB=AD$。我们令杆的终点G，固定在由条件 $DG=AB$ 确定的某点上。那么 $ABGD$ 是一个菱形，因而它的对边 AB 和DG 在连杆的所有位置始终平行。也就是 AB 始终平行于固定直线 DG，因此始终平行于它本身。这样终于获得了直线运动。

8. 重新安排肯普的二重菱形,可以按另一特别巧妙的方法产生直线运动。我们如果有二个二重长菱形连杆 $ABCDEF$ 和 $A'B'CD'E'F'$，其中之一是另一个的镜像。我们连接这二个菱形，使它们在C相交，因此 DCE' 和 $D'CE$ 是直杆,即在C处并不弯曲(图 66)。如我们已看到的角DCE 的角平分线是 CZ，它平行于 AB。类似角 $D'CE'$ 的角平分线 CZ' 平行于 $A'B'$。由于所作 DCE 和 $D'CE'$ 是对顶角，二条平分线 CZ 和CZ' 在一条直线上。AB 和 $A'B'$ 平行于这条直线，那么它们在连杆的任一位置彼此都是平行的。现在将证明直线 $A'B'$ 总是在直线 AB 的延长线上。由于已知这二条直线是平行的，我们还需证明的是点C和直线 AB 和 $A'B'$ 的距离相等。为此，我们只要证明二个长菱形 $ABCD$ 和 $A'B'C'D'$对于连杆的任一位置总是全等的。首先，由于角 DCE 和 $D'CE'$ 是对顶角，它们是相等的。由第 7 节，这些角是每个长菱形的对边夹角的二倍。二个长菱

形的对应边是相等的。如果我们能证明，长菱形完全是由它的边长以及二边之间的角度确定的，那么如果这些量相等，

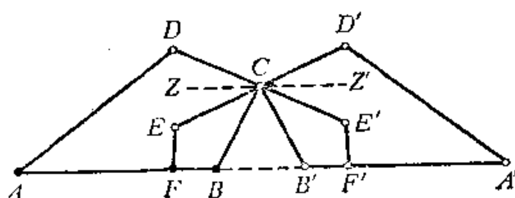


图 66

这二个长菱形将是全等的。为了检验长菱形已完全确定，我们在长菱形 $ABCD$ 内部作菱形 $BCDH$（图 67）。点 A、H、C 由于对称性，处在一条直线上。这个图形恰好是皮塞勒网。由于 HB 平行于 DC，那么角 ABH 等于长菱形对边所夹的角。它因而是图 66 中已知角 DCE 的一半。由于 AB 是给定的，且 BH 是等于给定的 BC，则三角形 ABH 完全确定。进而 D 作为 B 对于 AH 的镜像而被确定出来，而 C 由 B,H,D 所规定的菱形 $BCDH$ 的第四个顶点 C 也就定出。

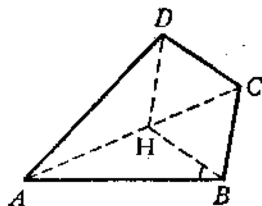


图 67

所以，我们可以断定，图 66 的二部分彼此总是全等的。因而 C 在 AB 上面的高度，等于它在 $A'B'$ 上面的高度。这就表示，AB 和 $A'B'$ 是在同一条直线上。

现在，如果我们使 A 和 B 固定，而杆 $A'B'$ 可以运动，但必须处在 AB 线上。这样我们从这个连杆获得了一个直线

运动。

现在考察的连杆较以前的复杂些。在先前的装置中，是一个点获得直线运动，而在这里是指整个杆 $A'B'$ 在它所处的线上运动。由于任一图形都能牢牢附在 $A'B'$ 上，这个装置可以完成任一图形或甚至整个平面上的平行移动，这样所有的点沿着相等且平行的线段运动。

19. 完 全 数

欧几里德的《原本》一书的第Ⅸ卷，是论述算术的第三册也是最后一册书。这部书包括我们在第一篇中已介绍过的素数无限性的证明，而且还包含有所谓完全数的讨论。柏拉图也提到过完全数，特别是在他的《共和国》一书的一个隐晦的段落中，在一次关于优生学的含糊不清的讨论中，他介绍了“姻数”。

完全数的题目，以及后来证明的有关它们的定理，是近代数学中为数不多的有趣的珍品。但我们将要讨论这个小论题，因为在欧几里德所采用的方法中，燃点起了下一篇我们将见到的那种思想的火花。而欧拉再次把它点燃，使之突然变成巨大的火焰：素数分布的近代理论。

欧几里德定义的完全数是这样一种数，即它等于它的所有因子的和。例如，6是一个完全数，因为它的因子通过试验容易知道是1,2,3,而 $1+2+3=6$。在这个完全数的定义中，数本身显然不能当作它的一个因子。通过不断地试验，求得6以后的下一个完全数是 $28=1+2+4+7+14$。下一

个完全数相当大，不易通过试验求得，但这无关紧要。我们真正需要的是，能找到完全数的一种一般的方法。

1．很明显，素数不可能是完全数。1和P是素数P的所有因子，而P本身还不能当作因子。因此所有因子的和就是1，但1当然不等于P。

2．除了注意到这个极简单的情形，我们可以考察一种略为复杂的情形。数9不是素数，但它是一个素数的平方。试验由1到8的数，我们看到1和3是可以考虑的所有因子。由于$1+3=4$，小于9，所以9不是完全数。

相应地，数1和P显然是素数P的平方P^2的因子。现在我们不能象9的情形那样，通过试验来检查并列举出所有因子。然而，我们可以**证明**，这二个因子就是所有可能的因子。欧几里德在建立了必要的引理后，作出了这个证明。除了公式表示的方法以外，我们的证明和欧几里德的基本相同。这个证明的基础是第11篇所证明的素因子分解唯一性定理。如果P^2有另外任一个因子，就可以用某种方法得到与$P^2=P\cdot P$不同的素因子分解。事实上，这另一种素因子分解必有不同于P的另一素因子，而由于素因子分解的唯一性定理，这是不可能的。

3．一般来说，我们可以用同样的方法证明，素数的任何次幂都不是完全数。正如P^2那样，我们知道P^2的可以考虑的所有因子是

$$1, P, P^2, \cdots P^{a-1}. \qquad (1)$$

由于它们组成几何级数，它们的和由一个著名的公式给出。这个公式是欧几里德为当时的目的而证明的。事实上

$$1+p+p^2+\cdots+p^{a-1}=\frac{p^a-1}{p-1}. \qquad (1a)$$

由于对于最小的素数 $p=2$ 来说，分母 $p-1$ 是 1，而对于更大的因子，分母变大，(1a)中的分数就决不大于它的分子 p^a-1。因此 p^a 的所有因子的和决不大于 p^a-1。它小于 p^a，因而 p^a 不是完全数。

4．现在我们看怎样来讨论象 $72=2^3\cdot 3^2$ 这种含有 2 个素数的数。我们直接来考虑 p^3q^2。p^3q^2 的因子，包括数 p^3q^2本身，是

$$\begin{array}{c} 1, p, p^2, p^3, \\ q, qp, qp^2, qp^3, \\ q^2, q^2p, q^2p^2, q^2p^3, \end{array} \qquad (2)$$

也就是所有形如 $p^\alpha q^\beta$ 的数，其中 α 最多大到 3，β大到 2。显然，所有这些数都可以整除 p^3q^2。再一次由素因子分解的唯一性定理得知，这些是所有因子。表(2)使我们能完全观察到这个数的全部因子。为了求因子的和，我们认为第二行就是第一行乘以 q，第三行是第一行乘以 q^2。第一行是几何级数 $1+p+p^2+p^3$ 的和，那么整个和是这个和的 1 倍，q倍，q^2 倍的和。因而整个和是

$$D=(1+q+q^2)(1+p+p^2+p^3)。$$

类此，$N=p^aq^b$ 的因子和是

$$D=(1+q+q^2+\cdots+q^b)(1+p+p^2+\cdots+p^a)$$

$$=\frac{q^{b+1}-1}{q-1}\frac{p^{a+1}-1}{p-1},$$

必须记住，在这个公式中，数N本身是包括在因子中的。

5．上面最后的结果，可以推广到含有多于 2 个素因子

的数N。对于 $N=p^a q^b r^c$，可列出象(2)那样的表，然后把整个表乘以r和乘以r^2，直到最后乘以r^c。总起来，整个表乘以$(1+r+r^2+\cdots+r^c)$。如果N还有另一个素因子，那么因子的和，就对应于另外的部分。一般来说，对于$N=p^a q^b r^c\cdots$，N的因子的和是

$$D=(1+p+p^2+\cdots+p^a)(1+q+q^2+\cdots+q^b)(1+r+r^2+\cdots+r^c)\cdots$$
$$=\frac{p^{a+1}-1}{p-1}\frac{q^{b+1}-1}{q-1}\frac{r^{c+1}-1}{r-1}\cdots, \tag{3}$$

其中仍然包括数N本身。

6．在欧几里德的著作里，基本上包含了以上这些内容，但只阐明 $N=p\cdot 2^b$ 的情形。在这种情形下，N是二个因子的幂的乘积，一个是2的幂，另一个是素数P的一次幂，公式(2 a)成为

$$D=\frac{2^{b+1}-1}{2-1}\cdot\frac{p^2-1}{p-1}.$$

如果N是完全数，就有 $D=2N$，因为N本身包括在因子中。并且

$$D=(2^{b+1}-1)(p+1)=2N.$$

代入N的值，得出

$$(2^{b+1}-1)(p+1)=2\cdot p\cdot 2^b=2^{b+1}\cdot p,$$
$$2^{b+1}(p+1)-(p+1)=2^{b+1}p,$$
$$2^{b+1}=p+1,$$
$$p=2^{b+1}-1. \tag{4}$$

因此，如果P不仅是素数，而且还等于$2^{b+1}-1$，那么N就是完全数。数$2^{b+1}-1$对于任意的b值不总是素数，但是我

们有欧几里德定理：

若 $2^{n+1}-1$ 是素数，则 $N=(2^{n+1}-1)\cdot 2^n$ 必是一个完全数。

7．为了求得更多的完全数，我们可以试验 $n=1,2,3,\cdots$ 的值：

$$n=1,\ N=(2^2-1)\cdot 2=3\cdot 2=6,$$

$$n=2,\ N=(2^3-1)\cdot 2^2=7\cdot 4=28,$$

$$n=3,\ N=(2^4-1)\cdot 2^3=15\cdot 8$$

不是完全数，15不是素数。

$$n=4,\ N=(2^5-1)\cdot 2^4=31\cdot 16=496,$$

$$n=5,\ N=(2^6-1)\cdot 2^5=63\cdot 32$$

不是完全数，63不是素数。

$$n=6,\ N=(2^7-1)\cdot 2^6=127\cdot 64=8128,$$

$$\cdots\cdots$$

欧几里德的定理立即引起一个新问题：当 n 是什么值时，$2^{n+1}-1$ 是素数？向前迈一步，答案就立即可得。如果 $n+1$ 是一个合成数，设 $n+1=uv$，则 $2^{n+1}-1=2^{uv}-1=(2^u)^v-1$。现在，由几何级数求和的公式，我们有

$$x^v-1=[x^{v-1}+x^{v-2}+\cdots+x+1][x-1],$$

并且令 $x=2^u$，得到

$$(2^u)^v-1=[(2^u)^{v-1}+\cdots+(2^u)+1][(2^u)-1]。$$

这是二个因子的乘积。因而除非 $n+1$ 本身是素数，$2^{n+1}-1$ 就不能是素数。上面的表继续作下去，7 后的下一个素数是 11，所以下一个试验的值必是 $n=10$。由于，数 $2^{11}-1=2047=23\cdot 89$ 不是素数，$(2^{11}-1)\cdot 2^{10}$ 就不是完全数，但 2047 的分解，在试验过程中是很值得注意的。已经发现往后若干

个数都是完全数，但还没有一个求所有完全数的完整的方法，甚至对于已经发现的完全数也是如此。另一个问题是，还没有回答完全数是否有无限多个或者是否有最后一个。

8．欧几里德的结果只得出这个定理，但有各种证据表明，当时知道的东西还要多。例如，詹姆布里查在没有进一步解释的情况下叙述了这样一件事，即除了欧几里德给出的那些以外，没有其它的偶数完全数。还不知道在古代它是否已被证明，或者如果证明了，又是怎样证明的。然而，欧拉给出了一个基于公式(3)的证明，而这个公式在欧几里德的著作中实质上已经得到。虽然这个证明对以后的论题并不是重要的，但我们仍将介绍这个有趣的证明。

设N是任意一个偶数。那么2是N的一个素因子，而且还表现为幂的形式，如2^n。N的剩下的其它素因子都是奇数，而且它们结合在一起表示某个奇数u。于是我们有$N=2^n\cdot u$。如果我们按照(3)组成N的因子的和D，第一部分是

$$\frac{2^{n+1}-1}{2-1}=2^{n+1}-1,$$

第二部分，即假设利用(3)求得u的因子的和为U。我们现在有

$$D=(2^{n+1}-1)U。\tag{5}$$

如果N是完全数，那么上式必须等于$2N$，因为N本身包括在因子中。因此，如果N是完全数，我们求

$$(2^{n+1}-1)U=2N=2\cdot 2^n u=2^{n+1}\cdot u,$$

$$2^{n+1}U-U=2^{n+1}u,$$

$$2^{n+1}(U-u)=U=(U-u)+u,$$

$$(2^{n+1}-1)(U-u)=u。\tag{6}$$

现在U是包括u本身在内的u的因子的和，那么$U-u$就是不包括u本身在内的u的因子的和。同样，(6)可推导出，如果N是完全数，$(U-u)$能整除u，即$(U-u)$是u的因子。如果$U-u$本身是一个因子，同时又是所有因子的和，那么它必是u的唯一因子。由于1一定是u的因子，它必然就是这个单独的因子$U-u$，而u一定是素数。那么(6)简化为$u=2^{n+1}-1$，我们有$N=2^n(2^{n+1}-1)$。这就是欧几里德的偶数完全数的形式。

9．在此我们遇到第二个问题：是否存在着奇数的完全数？这个问题也还没有解决。直到现在还没有人找到一个奇数完全数，看来这样的数是不存在的，但这还没有得到证明。

10．柏拉图《法规》一书第五卷中的第一节使我们想到，欧几里德还未能做到通晓他那时的所有知识。柏拉图规定，在一个新找到的城市里，地块数和土地占有者的人数，应选择得使这个数有尽可能多的因子，如5040有60－1个因子。他指出立法者必须充分熟悉算术，才能安排任意规模的城市。这是一个数的因子个数的问题。要解决它，我们可以利用关于因子的和的讨论中的一部分内容。例如，p^3q^2的全部因子都在表(2)中，而且我们可以逐行数出它们的个数，$3\cdot4=12$。这个数包括数p^3q^2本身，所以我们必须减去1，才得到真正的因子数12－1。对于$N=p^aq^b$，一般N的真因子的数目是

$$P=(a+1)(b+1)-1,$$

至于对完全一般的情形$N=p^aq^br^c\cdots$，它是

$$P=(a+1)(b+1)(c+1)\cdots-1。$$

特别是，对于$N=5040=2^4\cdot3^2\cdot5\cdot7$我们有

$P=(4+1)(2+1)(1+1)(1+1)-1=5\cdot3\cdot2\cdot2-1=60-1$。这就是柏拉图利用不寻常但并不周全的记法 60－1代替 59 的大概的解释；并且由这个记法，以及根据他关于立法者必须知道回答这个问题的建议，使得人们相信，他已经掌握了这个解法。

11．如果我们设想一个数的因子的个数与因子的和的问题是某个一般题目中的二种特殊情形，那么它们之间的密切联系会变得较为清楚，设数N的因子的s次幂的和是S。例如，对于 6，我们有 $S=1^s+2^s+3^s$。当 $s=1$,S 是N 的真因子的和 D。而当 $s=0$ 时，每一项都是 1，而 S 就是N 的真因子的个数。当 $s=2$，它是N的真因子平方的和，如此类推。对于 s 的每个值，用同样的论证可以找到类似于 (3) 的一个公式。值 $s=-1$ 可以和其它值一样地运用。这个-1次幂是因子的倒数，那么，当$N=6$，S 是

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}。$$

类似于公式(3)，N的因子的倒数显然是

$$R=\left(1+\frac{1}{P}+\cdots+\frac{1}{P^a}\right)\left(1+\frac{1}{q}+\cdots+\frac{1}{q^b}\right)\left(1+\frac{1}{r}+\cdots+\frac{1}{r^c}\right)\cdots。\qquad(7)$$

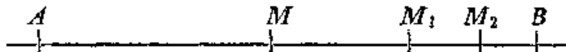
以上所有这些和，都是一个完整理论的一部分。这个理论从这一事实产生，即象(2)那样的表，能给出一个数的全部因子的完整清单。这个理论的整个基础，古代数学家已经掌握，这可在柏拉图那里得到证实，而且看来他们已经认识到它的美妙和重要。

20. 欧拉关于素数无限性的证明

我们在第一篇中讨论过的素数无限性的欧几里德证明，是在他认识完全数以前作的。欧拉加深和扩展了对完全数的研究，给出了素数无限性的另一个证明。这个证明所利用的思想，正是完全数理论的基础。

在进行证明以前，必须先作二个简单的观察。

1．设 AB 是只有 2 英尺的直线段。



如果我们由 A 走到直线段的中点 M，再由这里走到余下的那段 MB 的中点 M_1，由 M_1 再到又剩余的 M_1B 的中点 M_2，如此类推，我们总是与 B 相距一小段，但却是不断的接近 B。我们走过的距离应是 1 英尺，$\frac{1}{2}$英尺，$\frac{1}{4}$英尺，如此类推。所以这些距离加在一起应小于 2 英尺：

$$1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^n}<2.$$

如果线段不是按$\frac{1}{2}$，而是按另一个比 $x<1$ 不断地减少，同样的设想也是对的。要证明这个，我们将再一次利用几何级数求和的公式，

$$1+x+x^2+\cdots+x^n=\frac{1-x^{n+1}}{1-x}=\frac{1}{1-x}-\frac{x^{n+1}}{1-x}.$$

由于 x 小于 1，右端的第二个分数是正数，它是第一个数减

去的数，因此我们有

$$1+x+x^2+\cdots+x^n<\frac{1}{1-x}.$$

如果 p 是任意一个素数，则有 $\frac{1}{p}<1$，并且可以在这个不等式中用 $\frac{1}{p}$ 代换 x。我们可以断言，如果 p 是任意一个素数，并且 n 是整数，则

$$1+\frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^n}<\frac{1}{1-\frac{1}{p}}=\frac{p}{p-1}. \tag{1}$$

2. 如果我们写出

$$A_m=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{2^m}.$$

那么

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}>1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)=1+\frac{2}{2},$$

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}>A_2+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)>1+\frac{3}{2},$$

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{16}>A_3+\left(\frac{1}{16}+\cdots+\frac{1}{16}\right)>1+\frac{4}{2},$$

并且一般有

$$A_m=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{2^m}>1+\frac{m}{2}. \tag{2}$$

如 m 增大，则 A_m 也增大，并且只要 m 充分大，A_m 就能大

于我们事先给定的任意数。例如，为了使$\frac{m}{2}=999$，从而使$A_m>1000$，我们只需要选择$m=1998$就可以了。按照同样的方法，我们可以选择足够大的m，以使A_m大于1百万[1]。

3．设p是一个素数，并且写出(1)关于$2,3,5\cdots$直到p的每个素数的式子：

$$1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^n}<\frac{2}{2-1},$$

$$1+\frac{1}{3}+\frac{1}{3^2}+\cdots+\frac{1}{3^n}<\frac{3}{3-1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$1+\frac{1}{p}+\frac{1}{p^2}+\cdots+\frac{1}{p^n}<\frac{p}{p-1}.$$

现在我们作所有这些和的乘积R_p。它应小于所有右端的乘积M_p。

$$R_p<M_p=\frac{2}{1}\cdot\frac{3}{2}\cdot\frac{5}{4}\cdot\frac{7}{6}\cdots\frac{p}{p-1}.$$

由前一篇我们已熟悉了R_p这一类的乘积。当遍乘后，R_p是所有

$$\frac{1}{2^\alpha}\cdot\frac{1}{3^\beta}\cdots\frac{1}{p^\mu}$$

这些项的和。其中$\alpha,\beta,\cdots,\mu$的每一个，可取直到n的所有值。换句话说，R_p是

$$N=2^n\cdot 3^n\cdot 5^n\cdots p^n$$

的所有因子的倒数的和。因此R_p是整数的倒数的和，但不是全部可能的整数的倒数和。这些数只是由素因子2，3，5，$\cdots$，p以及它们的不高于n次的幂组成的。例如

$$\left(1+\frac{1}{2}+\frac{1}{2^2}\right)\left(1+\frac{1}{3}+\frac{1}{3^2}\right)=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{6}+\frac{1}{9}+\frac{1}{12}+\frac{1}{18}+\frac{1}{36}.$$

是 $2^2\cdot3^2=36$ 的所有因子的倒数的和。

4．我们现在转到实际的证明。设 m 是任意正整数。那么，A_m 是所有如下的数

$$1,2,3,4,\cdots,2^m \tag{3}$$

的倒数的和。我们考察那些是(3)中某个数的素因子的全部素数，并且称这些素数中最大的一个为 q。这样，(3)中的数只是素因子 $2,3,5,\cdots,q$ 以及它们的乘积。而且，这些素数的幂不可能高于 m 次幂。因为如果由最小的素数 2 表示的数是高过 m 次幂的，那么这个数就大于 2^m，所以不在数列(3)中出现。对于素数 2 是对的，那么对于所有较大的素数也一定是对的。因此(3)中的数全部包含在 $2^m\cdot3^m\cdot5^m\cdots q^m$ 的因子之中。

现在，如果我们不是对任意的 p，而是对我们的特殊的素数 q 组成第 3 节中的乘积 R_p，并且用 m 代替 n，那么它应包括 A_m 的所有的项。式子 R_q 将含有与 A_m 不同的其它的项，但重要的是，Am 的所有的项在 R_q 中都可以找到。结合第 3 节末尾的例，我们看到 $A_2=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}$ 都包括进去了，并且还有其它很多项。

由于 A_m 和 R_q 之间的这个关系，我们有 $A_m<R_q$。而且，由第 3 节，我们有 $R_q<M_q$，同时由(2) $1+\frac{m}{2}<A_m$。把这些结合在一起，有

$$1+\frac{m}{2}<\frac{2}{1}\cdot\frac{3}{2}\cdot\frac{5}{4}\cdots\frac{q}{q-1}. \qquad (4)$$

现在 m 是一个任意正整数，而 q 是(3)中全部数中一切素因子中最大的一个。(4)的左端，只要 m 选择的足够大，就能任意大。假设只存在有限个素数，则 q 可以成为最后一个素数而不能再继续增加。那么(4)的右端将增加至某一个确定值，然后保持为常数。这就和左端可以任意大这个事实矛盾了；因而这就证明了素数的个数是无限的。

这个证明比第一篇中的证明复杂多了。但是它的重要意义在于，同样的方法可以应用在很多类似的更为困难的问题上，其中很少的一些问题已在第一篇中提到过。这个方法是素数分布理论的基础，而这个理论是近代数学中最广泛最困难的领域中的一个。至少在实质上，这里的基本思想在古代已经掌握了。

[注] 由此可见$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}$，当 n 增大时，是无限增加的（其中 n 不限制为 2 的幂，但当它增加时，最终必将超过 2 的任意确定的幂）。在数学中，这就用"无穷级数 $1+\frac{1}{2}+\frac{1}{3}+\cdots$是发散的"的说法表示，虽然其中单独的项可以不断地变小。但是，在现在的题目中，并不包括这个重要事实，也不包括收敛和无限过程的概念。

21．极大问题的基本原理

我们已经多次讨论过极大问题。在第 3、5 和 6 篇中，我们列举了一些有关极大问题的数学小品，这些小品是有巨大

才能的数学家在写作更为重要和完美的作品的同时写出的。在这一篇里，我们将讨论作为所有这类问题的基础的一些原理。

这些原理可以通过研究一个极简单的极大问题提出。假设有一个给定的三角形(最好设想它是由纸剪下的)。我们的问题是，求包括边界在内的三角形上相距最远的二点P和Q(图68)。答案是容易猜到的：P和Q是最大边的端点。但我们怎样证明它呢？

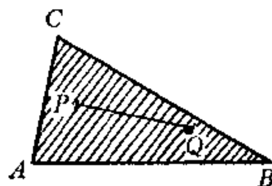


图 68

有一个在前几篇里不曾用过的简单方法。用它可以使我们这里的问题以及其它问题都得到解答。我们论述如下：如果有一点P是在三角形的内部，那么PQ一定不会是最大长度。因为在PQ的延长线上必有一点P_1，它和Q的距离大

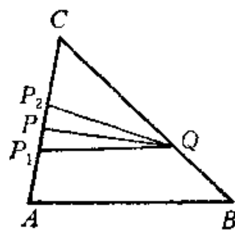


图 69 a

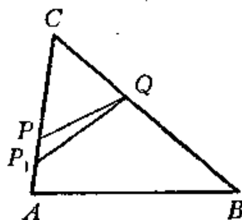


图 69 b

于P和Q的距离，并且它仍然在这个三角形的内部。如果P和Q都在三角形边界上，但其中之一，设为P，不是顶点，那么我们可以在边界上找到与P邻近的一点P_1，它和Q的

距离大于 PQ 的距离。PQ 垂直于 P 所在的边，以及 PQ 不垂直于这个边，(图 69 a，69 b) 这个结论都是明显的。因此，只有在 P 和 Q 都是端点时，PQ 才能是最大值；否则它就一定不是。由于 PQ 是三角形的一边，当然必定是最长的一边。

对于多边形，同样的论证使我们得到相应的结果：**多边形上相距最远的二点，必是顶点中相距最远的二点**。这个结果并不需要限定多边形是凸的。在图 70 的四边形中，底部二个端点是相距最远的点。

前几篇中的极大问题可以用同样的方法处理。使用这个原理求结果，要比用那几篇里的方法简捷。例如，在已知圆内求最大的内接三角形(图 71)，我们这里假定 ACB 不是等边的。如果 AC 和 BC 不相等，我们在圆内作以 AB 为底的



图 70

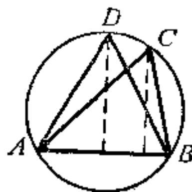


图 71

等腰三角形 ABD。这新三角形的高线比原来的长，因此有较大的面积。所以不等边的三角形不可能是最大的；最大的三角形必定是等边的。同样的方法也可用于垂足三角形的情形。

在前几篇中，我们为什么要用相当长而又比较复杂的证明呢？为什么不满足于这个简单得多的方法呢？理由是，这

个方法含有一个严重的逻辑错误。这个错误在二个世纪内都一直未被指出，直到十九世纪下半叶维尔斯特拉斯才把它揭露出来。

如果我们把原理应用于另一个例子(图 72)，这个错误将变得明显起来。由于平面无限延伸，显然在这个图形里没

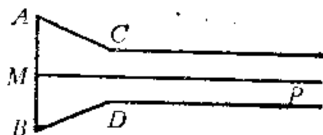


图 72

有相距最远的二个点；点 P 向右移得越远，距离 MP 也就越大。然而前面的论证照样可以进行。假设 P 和 Q 是在图形内部或边界的任意处，甚至它们可能是四个顶点 A, B, C, D 中的任一点。除非 PQ 恰好是边 AB，那么总能找到邻近的点 P_1，使 PQ 的距离增加。正如前面的情形，除了点对 A，B 外，对于任意的点对 P, Q，我们都能在它们邻近找到相距更远的点对。也就是除 A, B 外，没有其它点对可以给出最大值。如果现在严格的按照前面的论证进行，我们必须断定 AB 是最大值。

在这个最后的论证中，我们遵循完全一样的原理，得出了一个明显错误的结果。对于三角形，我们怎能保证我们得出的结果不错呢？

这个方法中所含的错误是这样的：对于三角形来说，我们最终证明了的是，除了最大边的端点外，没有其它的点对可以有最大的值。但是，这并没有告诉我们，这二个点比三

角形内其它任何点对相距更远。如果我们知道最大值问题必有一个解，我们就能合乎逻辑地断定最大值必是这个点对，因为没有其它的可能性。但是我们必须知道存在着一个解。如图 72 的情形，如果存在一个解，它应当是 AB，但在这个问题中，最大值实际上是不存在的。

现在很清楚了，为什么在前几篇中我们不能避开看来很麻烦的证明。这些证明不只是为使推理完善所希望的，并且是为避免一个严重的逻辑错误所必需的。为了完成本章的所有题目，我们还需要用对图 68 所作的有所欠缺的论证。

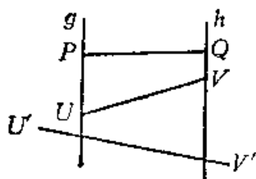


图 73

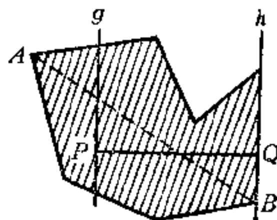


图 74

我们首先作一个简单的观察，如果我们有二条平行直线 g 和 h(图 73)，那么垂线 PQ 比二点分别在不同直线上的任意其它线段 UV 都短。也就是 PQ 比连接左边 g 上的一点 U' 和右边 h 上的一点 V' 的任意线段$U'V'$ 都短。

不限于对三角形，我们可以容易地对一般的多边形进行证明。

令 P 和 Q 是多边形内部或边界上的任意二点。我们在 PQ 的端点作二条垂直线 g, h(图 74)。这二条直线把整个图形分割出一条平行带。现在，由于多边形内至少有这个点 P

在 g 上，多边形的某些端点 A 必定位于 g 的左边，或可能在 g 上。而另外有某些端点 B 必定位于 h 的右边，或可能在 h 上。AB 距离不能小于 PQ，因而 PQ 不大于多边形任意二个端点间的最大距离。因此，这最后的距离就是最大值。

在下一篇里，我们将提出另一个极大问题，它比前面的任一问题更复杂。这个问题是，周长已知，求由直线或曲线所围成的面积最大的图形。我们将证明，最大值只可能是圆，但不打算证明圆的面积实际大过其它任何图形。由于曲线图形进入研究的范围，这个问题就复杂化了。因此，即使为了得到这有限的结果，我们也势必要应用一系列新的概念和事实。

22. 一定周长下面积最大的图形

为什么肥皂泡是球形的？这是因为由一定物质组成的肥皂膜内有一种内聚力，这种内聚力使肥皂膜变厚而面积缩小。当空气的压力不起作用时，被包围的空气将保持一固定的体积，而其表面积则将变得尽可能地小。换句话说，肥皂泡解决了在一定体积下求表面积最小的立体图形的问题。

我们现在要解决的问题，比任何肥皂泡和水滴所解决的问题简单些。为了代替考察立体图形，我们可把自己局限在二维的图形上，并要求作出既定面积下最小周长的图形。

1．这里我们要求的是既定面积下最小周长的图形，而本篇的标题要求的是一定周长下面积最大的图形。但它们之

间的差别只是表面的。实际上，假定已知周长为 k 的图形不是圆 K(图 75)，而已知周长为 f 的图形 F 有最大面积。那么我们有 $f=k$ 和 $F>K$。现在我们可以使图形 F 按比例缩



图 75

小，为此，只要把由图形 F 内部任意一点 O 到周界的所有射线 $OA, OB, OC, OD, \cdots$ 都按一定的比例缩减就可以了。我们可以选择这样一个比例，使缩减后的周长为 f' 的图形 F'，与 K 有同样的面积。因为显然图形周长同时减小，即 $f'<k$。因此，F' 是和圆有相同面积但周长却较小的一个图形。如果证明这是不可能的，那么标题指出的问题也就解决了，并且这个解就是圆。由于按相反的次序用同样方法也能加以证明，所以二个提法是完全等价的。

2．我们刚才已经以隐蔽的方式应用了下述关于图形的面积和周长的定理：

Ⅰ．如果平面图形按 $1:r$ 的比例，以 O 为相似中心相似缩减，那么周长以同样的比例 $1:r$ 减小，而面积以 $1:r^2$ 的比例减小。

我们还将应用其他几个有关图形面积和周长的定理。但我们将不提它们如何证明。因为这将要求对所含的全部概念的确切含义作一分析，而且还必须建立一套完整的理论。这样做并不是我们的目的，但我们将在此列出那些将要应用而

未加证明的所有定理。

Ⅱ．如果一个图形的表面是另一个图形的表面的一部分，那么这个图形的面积就小于另一个图形的面积（图 76）。

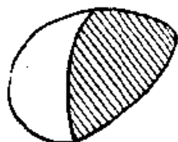


图 76

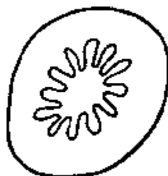


图 77



图 78



图 79

关于图形周长的类似论述是不能成立的。这可以从图 77 看出。这里我们只要使内部的曲线充分弯曲，就会使它大于外部的曲线。但是，如果只限制在凸图形，这个定理还是正确的。

如果一个图形的任意二点都可以用全部在图形内的一条直线连接起来，则该图形称为凸图形。圆和三角形是凸图形，而图 78 则不是，因为 P 和 Q 是在它内部，但连接它们的直线并不全部在图形内。

Ⅲ．如果一个凸曲线包围了另一个凸曲线[1]，则包围曲线的周长较大（图 76）。

Ⅳ．任意一个非凸图形都可以把它填补成面积增大周长减小的一个凸图形(图 79)。

未加证明而被接受的这四个定理，在直观上是十分明显的。如果有某些怀疑的话，那只是定理Ⅲ。然而，我们知道，直观验证并不是那样重要的。

3．现在我们准备着手作这个证明，并采用斯坦纳所用过的论述。我们首先要证明，一定周长有最大面积的图形必定是凸的。这可以从Ⅳ和Ⅰ直接导出。如果按照Ⅳ，图形可填补成为凸的，那么周长减小，面积增大。如果再按适当的比例扩大所得到的图形，可使其周长和原来图形的周长一样；那么根据Ⅰ，面积还要增大。这就说明，对于每个非凸的图形，能作一个有同样周长但面积较大的凸图形。因此，非凸图形在一定周长时，无论怎样都不能有最大面积。

4．因此我们现在只需考虑凸图形。下面证明，对于一定周长的任意的凸图形，存在一个同样周长、面积至少相等而有一对称轴的凸图形。设 P 为曲线上任一点(图 80)；且令 Q 是曲线上那样一个点，它由 P 沿曲线度量的距离恰好是周长的一半。弦 PQ 把图形切割为二部分。如果它们有不同面积，我们选择面积较大的那部分，并使之反射在直线 PQ 上；如果它们面积相同，我们任选其一，并使之反射在 PQ 上。这一部分和它的镜面反射部分组成一个图形，其周长和原来的一样，而面积至少一样大。并且这个图形有对称轴。如果这个图形不是凸的，那么应用Ⅳ和Ⅰ，按第 3 节那样就可作成。但这只是增加了面积，而显然不会破坏对称性。

这是命题完整的证明，但还需指出另外一点。如果原来的图形是一个圆，那么 PQ 是它的一个直径，而经反射得到

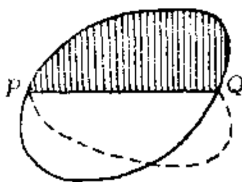


图 80



图 81

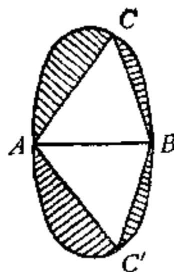


图 82

的图形仍是圆。如果原来的图形不是圆，那么被 PQ 分开的二部分至少有一部分不是半圆(图 81)。如果这部分的面积比那部份的大，那么经反射所得的新图形的面积较大。如果二部分面积相等，那么我们约定选择非半圆的那部分。所以新图形将不是一个圆。如果新图形不是凸的，我们可以利用Ⅳ和Ⅰ把它填补成凸的，再一次增加了它的面积。因此，我们可以说，或者有一个周长相等而面积较大的图形，或者这新图形是周长和面积都与原图相等的凸图形。再者，只有在原图形是圆的情形下，新图形才是圆。

5．如果原来的图形不是圆，我们可以“改善”它或用一个面积和周长都与其相等的凸图形来代替它。新图形有对称轴但不是一个圆。如果图形能被“改善”，它一定不是最大的。

我们现在要指出，如果具有对称轴的凸图形不是圆，我们可以作一个周长相等而面积一定增大的另一凸图形。如果这一点得到证明，那么就说明，不是圆的一切图形，都能“改善”。

如果 AB 是对称轴，(图 82)，那么，在图形的边界线上，必有一点 C，使角 $\angle ACB$ 不是直角。因为如果任一点所作的角是直角，那么根据熟知的一条定理（见图 18)，这条曲线应当是直径为 AB 的圆周。但按假设，这个曲线并不是圆周。我们现在另作一个直角三角形 $A_1B_1C_1$ (图 83)，使直角边 $A_1C_1=AC$ 和 $B_1C_1=BC$；这个三角形对 A_1B_1 的反射得到四边形 $A_1C_1B_1C_1'$ 的四条边。在这个四边形的边上，再放上原图由 $ACBC'$ 切割开的对应的扇形。(这些扇形在图 82、83 上画上阴影)

这样，所得到的新图形的周长与旧的一样，因为它由相同的四条曲线弧组成。新图形的面积，由四个扇形以及四边形 $A_1C_1B_1C'$ 组成。扇形在二个图形上是相同的，所以只需要比较四边形或甚至只比较相应的三角形 ABC 和 $A_1B_1C_1$，因为两个图形是对称的。三角形 $A_1B_1C_1$ 的面积为 $\frac{1}{2}A_1C_1\cdot B_1C_1$。三角形 ABC 由 B 到 AC 的高小于 BC，因为 C 点的角不是直角。所以，这个三角形的面积将小于 $\frac{1}{2}AC\cdot BC$。但 $A_1C_1=AC$，$B_1C_1=BC$；因此 ABC 的面积小于 $A_1B_1C_1$ 的面积。这就是说，新图形较原来的图形有较大的面积。

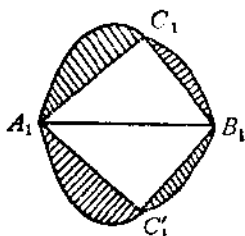


图 83

至此，证明已全部完成：对于不是圆的任意图形，可以作另一个周长相等而有较大面积的图形。因此，可能具有最大面积的图形只能是圆。

我们还未证明圆的面积，事实上已超过周长相等的其它图形的面积。我们已证明的部分与整个问题之间的关系，在前一篇里都已详细谈过。这个问题的彻底解决是可以办到的，但需要系统地建立起完整的理论，而这已不是本书的目的了。

[注] 凸曲线是指凸图形的边界线。——译注

23. 循环小数[1]

1. 分数化为小数是我们熟悉的过程。当这个过程完成的时候，小数会有各种不同的形式，试看下面的例子：

Ⅰ. $\frac{1}{5}=0.2$, $\frac{3}{40}=0.075$;

Ⅱ. $\frac{4}{9}=0.4444\cdots$, $\frac{1}{7}=0.142857142857\cdots$;

Ⅲ. $\frac{1}{6}=0.1666\cdots$, $\frac{7}{30}=0.2333\cdots$.

最简单的是第Ⅰ类。根据小数的意义，我们可以把它们写成以 10 的幂为分母的普通分数：

$$\frac{1}{5}=\frac{2}{10},\quad \frac{3}{40}=\frac{75}{1000}.$$

这些等式不能说明任何特别的东西。它们只能说明，已知的分数可以“扩大”，使得它们的分母成为 10 的幂的形式。这对于那些分母含有 10 的幂的因子的分数，都可以作到。这样一个分母的特点是，除了 2 和 5 以外，它没有任何其它素因子，因为如果 γ 是 α 和 β 中的较大者，数 $2^{\alpha}\cdot 5^{\beta}$ 可以整除

10^{γ}。所以，分母是 $2^{\alpha}\cdot5^{\beta}$ 的一个分数总是可以"扩大"成分母是 10^{γ} 的分数，并且最终变为 γ 位小数。

2．如果已知的分数的分母含有不能被 2 或 5 整除的因子 k，那么这种分数就不能"扩大"成分母是 10 的幂的分数。事实上，假设

$$\frac{1}{2^{\alpha}\cdot5^{\beta}\cdot k}=\frac{a}{10^{\delta}}=\frac{a}{2^{\delta}\cdot5^{\delta}},$$

那么应导出

$$2^{\delta-\alpha}5^{\delta-\beta}=a\cdot k$$

数 k 假设大于 1，它应该能被 2 或 5 整除，因为由 $a\cdot k$ 的素因子分解的唯一性，2 或 5 是能整除 $a\cdot k$ 和 k 的仅有的素因子。

Ⅱ、Ⅲ都是属于这一类分数的例子。在这种情形下，我们说，分数 $\frac{4}{9}$，$\frac{1}{7}$，$\frac{1}{6}$ 和 $\frac{7}{30}$ 可以展开为"无限小数"。但是，就其实质来说，这种小数是和分母是 10 的幂的那种意义下的小数完全不同的。因为小数展开中没有最后的数字，那么也就不存在确定的 10 的幂了。

当我们谈到"无限小数"时，我们用的"小数"这个词具有新的广泛的含义。就我们的目的来说，不需要强调这个新含义的严格意义。所涉及的仅是这个问题的形式方面，在这种情形下，分数转化为小数的过程并未间断。它无穷尽地产生无限小数的越来越多的数字。比较多的无限小数是循环的，即从某一确定的地方开始，数字的序列就成为同一数字组的一再重复。例如，$\frac{1}{7}$ 一开头就是 142857 的重复。因为只有

十个数字，所以任意一个无限小数中的某个数字必定在以后会重复出现。但是，这并不表示所有这样的小数是循环的小数。

$$0.101001000100001\cdots,$$

是无限不循环小数的一个例子。其中只用到数字0和1，而且第 n 个1的后面恰好跟着 n 个零。

3．在Ⅰ和Ⅱ中引进了其分母除2和5外，还含有其它素因子的分数的例子。我们现在要指出，这样的分数往往导致循环小数。首先研究$\frac{1}{7}$这个例。重复运用普通的除法可得小数：

```
   0.142857···
7)1.000000
  7
  30
  28
   20
   14
    60
    56
     40
     35
      50
      49
       1
```

当余数1一经出现，除法的整个过程将和最初的时候一样重又开始。余数的序列1，3，2，6，4，5以及商数序列1，4，2，8，5，7将一再重复出现。循环小数将有循环节142857。因为我们在其它例子中也要进行这种除法，为了今后方便，可用一种缩写形式来表示这个过程。我们写出

$$\frac{1}{7}=\underset{1}{0.}\overline{\underset{3}{1}\underset{2}{4}\underset{6}{2}\underset{4}{8}\underset{5}{5}\underset{1}{7}}\cdots,$$

在这里我们记下了在每一步求得一个新商后的余数。无限小数的循环节通常用它上面的水平黑线指明。另一个例子是

$$\frac{3}{41}=\underset{3}{0.}\overline{\underset{30}{0}\ \underset{13}{7}\ \underset{7}{3}\ \underset{29}{1}\ \underset{3}{7}}\cdots,$$

这里当余数 3 出现时，除法过程就开始重复，因为整个过程是从 3 开始的。商数同时将从循环节的终点开始与以前重复。现在的情形就是从数字 7 开始，它是因为余数开始重复而引起的。

对任意分数$\frac{a}{b}$作除法时，我们必然会遇到早先出现过的一个余数。(分数的分子当作第一个余数）余数只可能是 1，2，3，$\cdots$，$b-1$。0 不包括在余数中，因为它的出现就表示除法已经结束，而且小数是有限的。但是，如果分母有不同于 2 或 5 的素因子时，我们知道，这是不可能的。现在由于余数只可能有 $b-1$ 个数字，那么在除法过程中，在第 b 步或更早些，就必然开始重复。所以$\frac{a}{b}$的小数展开式是循环的，并且它的循环节至多是 $b-1$ 位。对于$\frac{1}{7}$，正好达到循环节的这个最大长度；至于分数$\frac{3}{41}$，则没有达到。另一个有循环节最大长度的例子是：

$$\frac{1}{17}=\underset{1}{0.}\overline{\underset{10}{0}\ \underset{15}{5}\ \underset{14}{8}\ \underset{4}{8}\ \underset{6}{2}\ \underset{9}{3}\ \underset{5}{5}\ \underset{16}{2}\ \underset{7}{9}\ \underset{2}{4}\ \underset{3}{1}\ \underset{13}{1}\ \underset{11}{7}\ \underset{8}{6}\ \underset{12}{4}\ \underset{1}{7}}\cdots,$$

它的循环节长度等于 $b-1=17-1=16$。

4. 从现在起，我们将考察的仅是分母与 10 互素的分数，即分母不含素因子 2 与 5 的分数。我们将着重谈谈循环节的长度。

分数$\frac{a}{b}$当然是既约分数；a 与 b 是互素的。于是相除得到的余数都将是与 b 互素的数。设 r 是与 b 互素的余数。除法的下一步是由 b 除 $10r$ 得到商 q 和余数 r_1。这就是说，$10r$ 含 b 的 q 倍，并多出余数 r_1：

$$10r=qb+r_1,$$

或

$$r_1=10r-qb. \tag{1}$$

因为 b 与 10 互素，也与 r 互素，于是 b 既与 $10r$，也与 $10r-qb$ 互素。因此，对于每一个与 b 互素的余数 r 都随之有与 b 互素的另一个余数 r_1。由于分子 a 被当作第一个余数，并且与 b 互素，因此以后产生的所有余数都与 b 互素。由此得到结论，**小数$\frac{a}{b}$的循环节的长度不大于与 b 互素的余数的个数。**

与 b 互素的余数的个数本身是很有趣的数。在数论中，它常用符号 $\varphi(b)$表示。例如，$\varphi(2)=1$，$\varphi(3)=2$，$\varphi(4)=2$，$\varphi(5)=4$，$\varphi(6)=2$，$\varphi(7)=6$，$\varphi(8)=4$。特别是对于素数 p，因为它前面所有的 $p-1$ 个数都与它互素，我们有

$$\varphi(p)=p-1. \tag{2}$$

利用符号 $\varphi(b)$，我们现在可以说：当 b 与 10 互素时，分数

$\frac{a}{b}$的循环节的长度，最多有$\varphi(b)$位。

5．我们已经看到，如果b有不同于2和5的素因子，甚至当b与10不互素时，分数$\frac{a}{b}$的小数展开式必定是循环的。我们看到有限个可能的余数中必然总会有一个重新出现。据此，我们可以断定它必定是循环的，但还不能肯定循环将从何处开始。在例Ⅰ中，循环从小数点后立即开始，而在例Ⅱ中，只是在小数点后经过1个数字后开始。在例Ⅰ中，分母与10没有公共因子，而且我们将证明，若b与10互素，则$\frac{a}{b}$的小数展开式必从小数点后立即开始循环。为此我们就得证明，开始重复的第一个余数是整个序列中的第一个，即分子自身。即如果二个余数相等：$r_m=r_n$，那么前面的余数r_{m-1}和r_{n-1}也相等，实际上可以是任意靠前的余数。由于$10r_{m-1}$和$10r_{n-1}$除以b得到r_m和r_n，

$$10r_{m-1}=q_{m-1}\cdot b+r_m,$$

$$10r_{n-1}=q_{n-1}\cdot b+r_n,$$

应用$r_m=r_n$，由减法求得

$$10(r_{m-1}-r_{n-1})=(q_{m-1}-q_{n-1})b,$$

由此可见$10(r_{m-1}-r_{n-1})$能被b整除。由于b与10互素，则b必能整除$r_{m-1}-r_{n-1}$，那么这个差$r_{m-1}-r_{n-1}$必是下列数中的一个，

$$0,\ \pm b,\ \pm 2b,\ \pm 3b,\ \cdots,$$

现在，另一方面r_{m-1}和r_{n-1}是余数，每一个都小于b，因此这个差的数值小于b。仅有的可能性是

$$r_{m-1}-r_{n-1}=0,$$

$$r_{m-1}=r_{n-1}.$$

因此，循环应尽可能提前开始，实际上应从小数点后立即开始。

6. 如果用 λ 表示循环节的长度，那么当 $\frac{a}{b}$ 展开为小数时，在除法过程中经 λ 步后，第一个余数将重新出现。除法每进行一步，小数就向右移动一位。λ 步后出现的余数 a 实际上可用 $\frac{a}{10^\lambda}$ 表示。因此，如果用 b 除 $10^\lambda a$ 得余数 a，那么 $10^\lambda a-a$ 可被 b 整除，由于

$$10^\lambda a-a=a(10^\lambda-1),$$

以及 a 与 b 互素，这就表明，b 能整除 $10^\lambda-1$。由于循环节终止于余数 a 第一次重新出现的时候，所以 λ 是使 $10^\lambda-1$ 能被 b 整除的最小的数。也就是使 $a\cdot10^\lambda$ 除以 b 有余数 a 的最小的数。

我们已经证明

定理Ⅰ. $\frac{a}{b}$的循环节的长度λ是使 $10^\lambda-1$ 能被 b 整除的最小的数。

应该强调有二个事实包含在这个定理中。第一，每一个与 10 互素的数 b，对应一个使得 $10^\lambda-1$ 能被 b 整除的数 λ。这样的一个数 λ 的存在性不是自明的，但是作为循环节的长度，我们已经求得它了，并且在第 3 节中证明了循环节的存在性。第二，数 λ 由 b 确定，并且与 a 无关。有相同分母 b 的一切既约分数 $\frac{a}{b}$，都有相同长度的循环节。我们将用 $\lambda=\lambda$

(b)的记号强调λ对于b的依赖性。

7. 我们现在将研究$\frac{a}{b}$的小数展开式，其中a有不同的值而b是相同的。我们已经求得

$$\frac{1}{7}=0.\overline{\underset{1}{1}\ \underset{3}{4}\ \underset{2}{2}\ \underset{6}{8}\ \underset{4}{5}\ \underset{5}{7}}\ \underset{1}{}\cdots,$$

类此得到

$$\frac{2}{7}=0.\overline{\underset{2}{2}\ \underset{6}{8}\ \underset{4}{5}\ \underset{5}{7}\ \underset{1}{1}\ \underset{3}{4}}\ \underset{2}{}\cdots,$$

$\frac{2}{7}$的循环节 285714，可以通过对$\frac{1}{7}$的循环节 142857 的数字进行轮换得到。如果记下余数 2，$\frac{2}{7}$的展式就由此开始，很明显 2 也出现在$\frac{1}{7}$的余数中。由此开始一直往后，二个展式必然逐步都是相同的。余数 3,4,5,6 也将出现在$\frac{1}{7}$的展式中。因为$\frac{1}{7}$的展式有长度为 6 的循环节，所以事实上关于 7 的 6 个余数必然出现。如果我们设想循环节的数字是通过一个轮换得到的，这个轮换的最后一个数字后面就是循环节的第一个数字，那么可以用下表排列出余数和商：

余　数	1	3	2	6	4	5
商	1	4	2	8	5	7

由此表，我们可以写出$\frac{6}{7}$以及其它类似的分数的小数展开式。例如，$\frac{6}{7}$的循环节由 8 开始，并且是 857142，所以我们有$\frac{6}{7}=0.\overline{8\ 5\ 7\ 1\ 4\ 2}$。

我们以前说的循环节都是对于商来说的。然而，如我们已看到的，余数也有同样形式和同样长度的循环节。

8．$\frac{a}{b}$**的循环节长度不能超过** $\varphi(b)$**，但实际上并不一定要这样长。**对于$\frac{1}{7}$和$\frac{1}{17}$，我们求得它们的循环节有最大长度，但是对于$\frac{3}{41}$，我们求得的是 $\lambda=5$，小于 $\varphi(41)=40$。进一步确定循环节的实际长度是困难的，在每一种情形下，它必须通过计算求得。但关于循环节的长度 $\lambda(b)$，我们可以比 $\lambda(b)\leqslant\varphi(b)$ 谈得更多些。

我们将以新例$\frac{1}{21}$作为讨论的出发点。通过试验，小于 21 的 20 个数中很易求得 $\psi(21)=12$。继而，通过除法我们得到

$$\frac{1}{21}=\underset{1}{0.}\overline{\underset{10}{0}\ \underset{16}{4}\ \underset{13}{7}\ \underset{4}{6}\ \underset{19}{1}\ \underset{1}{9}}\cdots,$$

循环节的长度 $\lambda(21)=6<12=\varphi(21)$。在作除法的过程中，12 个可能的余数中，只有 6 个出现。我们把它们排列在表中：

(A)

余 数	1	10	16	13	4	19
商 数	0	4	7	6	1	9

由此表我们可以写出

$$\frac{10}{21}=0.\overline{4\ 7\ 6\ 1\ 9\ 0}\cdots,\ \frac{4}{21}=0.\overline{1\ 9\ 0\ 4\ 7\ 6}\cdots.$$

但是，我们不能求得另一个分数$\frac{2}{21}$的展式。这个 分 数需要重新作除法，

$$\frac{2}{21}=0\underset{2}{.}\overline{\underset{20}{0}\ \underset{11}{9}\ \underset{5}{5}\ \underset{8}{2}\ \underset{17}{3}\ \underset{2}{8}}\cdots,$$

由此我们得到第二个表

(B)

余 数	2	20	11	5	8	17
商 数	0	9	5	2	3	8

在此表中，不只是出现了一个新余数 2，所有其它余数也都是新的。可以证明以前的旧余数在此表中不会出现。每一个余数完全确定了以后展开的全部过程，因而表(A) 的任一个余数都能得到(A) 的 6 位数字循环节的所有余数。于是，如果(B)含有(A)中出现的任一个余数，(B)就应包含(A)中的所有余数。由于(B)只含有 6 个余数， 它不可能再含任何其它的数；而这不符合(B)含有余数 2 的事实。二个表 (A) 和(B)合在一起，现在包括了所有的 $\varphi(21)=12$ 个余数，这12个数可以作为以 21 为分母的既约真分数的分子。

9. 上面对于以 21 为分母的分数作表(A)和表(B)中所包含的思想，可以用来得到有关循环节长度的某些一般结论。

如果$\frac{1}{b}$的展式中有最大长度 $\lambda(b)=\varphi(b)$，就存在如第 7 节中$\frac{1}{7}$那种情形的一个表。

然而，如果我们求得，例如关于 $b=21$ 有 $\lambda(b)<\varphi(b)$，那么只有 $\lambda(b)$ 个余数出现在$\frac{1}{b}$的展式中。利用这些数作表(A)，显然此表不含有所有可能的 $\varphi(b)$ 个余数。我们选择一个与 b 互素且不出现在表(A)中的余数 r，根据第 6 节，$\frac{r}{b}$展开为小数后的循环节的长度也是 $\lambda(b)$。我们利用新的余数和商作第二个表(B)。(B)中的余数都不同于(A)中的余数，因为(B)含有 r，以及由(A)的任一个余数出发都能得到(A)的所有 $\lambda(b)$ 个余数。

表(A)和表(B)合在一起含有 $2\lambda(b)$ 个不同的余数，它们都与 b 互素。或者这些数表示所有可能的余数，这时 $2\lambda(b)=\varphi(b)$；或者还剩有其它余数。如果 s 是不在(A)和(B)中的一个余数，我们展开$\frac{s}{b}$，并且得到另一个表(C)，它包含 $\lambda(b)$ 个既不含于(A)也不含于(B)的新余数。三个表合在一起，我们现在将有 $3\lambda(b)$ 个不同的且与 b 互素的余数。如果 $3\lambda(b)$ 等于 $\varphi(b)$，所有可能的余数就都包括无遗了。或者，我们重复这个过程，再组成新表，直到所有可能的余数 $\varphi(b)$ 都用到了为止。重要的是，每一次当一个新的

余数出现，就有$(\lambda-1)$个别的新余数包含在内。

如这些表最后共有K个，它们含有与b互素的所有的余数$\varphi(b)$。每一个表含有$\lambda(b)$个余数，并且没有一个余数出现二次。因此

$$\varphi(b)=K\cdot\lambda(b)。 \tag{3}$$

这就得到了定理：

循环节的长度$\lambda(b)$是$\varphi(b)$的因子[Ⅱ]。

如果p是素数，我们已经知道$\varphi(p)=p-1$。因此我们有一个特殊的结果，即如果p是素数，那么分数$\frac{a}{p}$的循环节的长度是$p-1$的因子。对以前举过的那些例子来说，我们得到$\lambda(3)=1$，$\lambda(7)=6$，$\lambda(17)=16$，$\lambda(41)=5$。

作为这个方法的又一个例子，我们来研究分母是39的分数的余数是怎样分布在各种循环节中的。我们这样开始

$$\frac{1}{39}=\underset{1}{0}.\overline{\underset{10}{0}\ \underset{22}{2}\ \underset{25}{5}\ \underset{16}{6}\ \underset{4}{4}\ \underset{1}{1}}\cdots, \tag{A}$$

由此很容易作出余数和商数表。问题是此表是否包含了与39互素的所有余数。显然余数2不见了，那么可以利用它构作下表：

$$\frac{2}{39}=\underset{2}{0}.\overline{\underset{20}{0}\ \underset{5}{5}\ \underset{11}{1}\ \underset{32}{2}\ \underset{8}{8}\ \underset{2}{2}}\cdots, \tag{B}$$

这就产生了6个新余数。和39互素而且不在(A)与(B)中的最小的余数是7。由此，使我们又得出一个新表：

$$\frac{7}{39}=\underset{7}{0}.\overline{\underset{31}{1}\ \underset{37}{7}\ \underset{19}{9}\ \underset{34}{4}\ \underset{28}{8}\ \underset{7}{7}}\cdots。 \tag{C}$$

现在14是和39互素且不包含在(A)，(B)和(C)的18个余

数中最小一个余数。我们利用它继续进行我们的计算过程：

$$\frac{14}{39}=0.\overline{\underset{23}{3}\ \underset{35}{5}\ \underset{38}{8}\ \underset{29}{9}\ \underset{17}{7}\ \underset{14}{4}}\cdots. \quad (D)$$

总计，我们现在有 24 个不同的且与 39 互素的余数。这就最后用尽了所有可能有的余数。事实上，在由 1 到 39 的数中，3 和 13 的倍数和 39 有公因子。3的倍数在这些数中占$\frac{1}{3}$，共有 13 个，而 13 的倍数又有 2 个，就是 13 和 26（39 已算在 3 的倍数中）。合起来，从 1 到 39，和 39 有公因子的数有15个，剩下 $39-15=24$ 个是和 39 互素的。因此我们有 $\varphi(39)=24$，这和 $\lambda(39)=6$ 是符合的。并且由于 $\varphi(39)=4\cdot\lambda(39)$，所以有 4 个表，$(A)$，$(B)$，$(C)$ 和 (D)。

10．我们得到的有关循环节长度的结果，可以用来得到一个重要的一般定理。我们首先需要如下的简单引理：

如果 X 和 K 是正整数，那么 X^K-1 可以被 $X-1$ 整除。

这个引理可由下述公式直接得到

$$(1+X+X^2+\cdots+X^{K-1})(X-1)=X^K-1,$$

它又可由等比级数求和公式得到，或直接由乘法得到。如果我们取值 $X=10^{\lambda(b)}$，那么引理断定 $10^{\lambda(b)}-1$ 可以整除 $10^{K\lambda(b)}-1$。我们选取 K 为满足（3）的特殊值，那么我们有

$$10^{K\lambda(b)}-1=10^{\varphi(b)}-1.$$

因此 $10^{\lambda(b)}-1$ 是 $10^{\varphi(b)}-1$ 的因子。但根据第 6 节，b 可以整除 $10^{\lambda(b)}-1$，因而 b 也是 $10^{\varphi(b)}-1$ 的因子。我们把这作为定理写出：

如果 b 与 10 互素，那么 $10^{\varphi(b)}-1$ 可被 b 整除。这个定理已经和小数没有什么关系了，因为 $\varphi(b)$ 有完全独立的意

义。由方程 (2) 我们可以得到这种特殊的情形：

如果 p 是不能被 10 整除的素数，那么 $10^{p-1}-1$ 可以被 p 整除。

在这些定理中，数 10 的出现是非本质的。它的出现，只是由于普通的数系是建立在数 10 的基础上的缘故。如果设想我们所利用的数系建立在另一个数 g 的基础上，我们就可以用"g 进位的小数"代替普通的小数。我们的全部讨论可以不加改变地重述一遍，并且得到一般的定理：

定理Ⅱ. 如果 b 和 g 互素，那么 $g^{\varphi(b)}-1$ 可被 b 整除。

定理Ⅲ. 如果 p 是不能被 g 整除的素数，那么 $g^{p-1}-1$ 可被 p 整除。

这里得到的定理，已经推广到远远超出小数这个特殊题目的范围。它是数论中的一个基本定理。定理Ⅲ这种特殊情形，是费马发现的，称作费马定理，定理Ⅱ是欧拉推广费马定理而得到的。

用几个例子来解释这些定理：

$p=5$：

$$2^{5-1}-1=15=3\cdot5,$$
$$3^{5-1}-1=80=16\cdot5,$$
$$4^{5-1}-1=255=51\cdot5;$$

$p=7$：

$$2^{7-1}-1=63=9\cdot7,$$
$$3^{7-1}-1=728=104\cdot7,$$
$$5^{7-1}-1=15624=2232\cdot7,$$
$$10^{7-1}-1=999999=142857\cdot7;$$

$b=6$，$\varphi(6)=2$：

$$5^2-1=24=4\cdot 6,$$

$$7^2-1=48=8\cdot 6;$$

$b=9$，$\varphi(9)=6$：

$$2^6-1=63=7\cdot 9,$$

$$4^6-1=4095=455\cdot 9,$$

$$5^6-1=15624=1736\cdot 9;$$

$b=10$，$\varphi(10)=4$：

$$3^4-1=80=8\cdot 10,$$

$$7^4-1=2400=240\cdot 10,$$

$$9^4-1=6560=656\cdot 10。$$

11．作了这段插曲以后，现在我们再回到小数方面来。我们已经知道，分母 b 和 10 是互素的既约分数$\frac{a}{b}$，可以展开为一个在小数点后立即循环的小数。

如果已知这样一个小数，我们可以由它求得分数。设小数有长度为 λ 的循环节。我们假设 P 是有 λ 位数字的普通的一个数。例如，$\frac{1}{7}=0.\overline{142857}\cdots$，我们有 $P=142857$（十四万二千八百五十七）。如果小数是$\frac{a}{b}$的展式，那么余数 a 在 b 除以 a 的 λ 步后重新出现。在 a 出现时得到的那部分商，恰好是一个循环节。因此，$a\cdot 10^\lambda-a$ 可以被 b 整除，并且我们有

$$a(10^\lambda-1)=b\cdot P,$$

或（4）
$$\frac{a}{b}=\frac{P}{10^\lambda-1}$$

右端的分数通常不是既约的形式，但在化简过程中可以得到

我们所希望的分数$\frac{a}{b}$。由于2和5可以整除10，它们就不能整除$10^{\lambda}-1$，它们也就不能整除$10^{\lambda}-1$的因子b。这就证明了，每一个由小数点后立即循环的小数，必然是由分母与10互素的分数$\frac{a}{b}$得来。我们早已知道反过来也是对的。小数点后立即循环的小数叫作“纯循环”小数。

12．上面我们研究的是纯循环小数。然而，第1节之Ⅲ中的例子并不是纯循环小数。在小数点和循环节之间有一个或几个数字的小数叫作“混合循环小数”。由于有限小数产生于分母只有素因子2和5的分数，而纯循环小数产生于分母既不含2也不含5的分数，它们都与混合循环小数不同，所以混合循环小数产生于这样的分数，其分母含有和10有公因子的数以及和10互素的数。我们仅只指出这第三种情形；因为它没有提供我们值得研究的东西。

13．在作了循环小数这些基本性质的讨论以后，我们将通过研究另一个特性来结束本篇。这个特性与其说它重要，不如说它较为有趣。$\frac{1}{7}$的循环节由6个数字142857组成。我们把它对半分开，然后相加，得到

$$142+857=999.$$

$\frac{1}{17}$的循环节是0588235294117647。把它对半分开，然后加起来，得到

$$05882352+94117647=99999999.$$

$\frac{1}{11}$的循环节是09，我们有

$$0+9=9.$$

我们将证明，如果循环节产生于分母是素数的分数$\frac{a}{p}$，并且由偶数个数码组成，那么这个循环节的两半之和总是一个完全由数码9组成的数。

如果循环节的长度λ是偶数，我们可记为$\lambda=2l$。还有，如果循环节P的两半是A和B，P是有λ个数字的数，而A和B都是有l个数字的数。留意到一个数中的数字的位置的意义，我们有

$$P=A\cdot 10^l+B.$$

我们知道，分数$\frac{a}{p}$可以由循环节P并借助(4)求得：

$$\frac{a}{p}=\frac{P}{10^\lambda-1}=\frac{A\cdot 10^l+B}{10^\lambda-1}. \tag{5}$$

由于$\lambda=2l$，我们有

$$10^\lambda-1=10^{2l}-1=(10^l-1)(10^l+1). \tag{6}$$

等式(5)说明分母p可以扩大为$10^\lambda-1$。这就是说p可以整除$10^\lambda-1$。这是可以由第6节求得的事实。如果p可以整除$10^\lambda-1=(10^l-1)(10^l+1)$，由于它是素数，它必定至少可整除二个因子中的一个。由于l小于λ，而λ是使$10^\lambda-1$可被p整除的最小的数，所以p不能整除10^l-1。最后p必须能整除10^l+1。由(5)和(6)，我们有

$$\frac{a}{p}=\frac{A\cdot 10^l+B}{(10^l-1)(10^l+1)}.$$

这可改写为

$$\frac{a(10^l+1)}{p}=\frac{A\cdot 10^l+B}{10^l-1}.$$

由于 p 可以整除 10^l+1，左端就是一个整数，因而右端也是一个整数。现在还可以写为

$$\frac{A\cdot 10^l+B}{10^l-1}=\frac{A(10^l-1)+A+B}{10^l-1}=A+\frac{A+B}{10^l-1},$$

而且由于 A 是一个整数，因此

$$\frac{A+B}{10^l-1}=h. \tag{7}$$

现在 A 是由 l 个数字组成的，并且当所有这些数字都是 9 时是最大的。l 个 9 组成的数是 10^l-1，因此 $A\leqslant 10^l-1$。同样的方法也有 $B\leqslant 10^l-1$，因此

$$A+B\leqslant 2\cdot(10^l-1).$$

这个不等式的等号是不能成立的，因为 $A+B=2\cdot(10^l-1)$，这就可推导出 A 和 B 同时有值 10^l-1。那么 A 和 B 只应当含有 9，并且 p 就是由 $2l$ 个相同的数码 9 组成的数。这是荒谬的，因为这时循环节只由一个数码 9 组成，而我们假定的是有偶数个数码。因此我们有

$$A+B<2\cdot(10^l-1). \tag{8}$$

现在把(7)改写为

$$A+B=h\cdot(10^l-1). \tag{9}$$

其中 h 是正整数。根据(8)，h 小于 2，那么它必然是 1。从而(9)变为

$$A+B=10^l-1,$$

即 $A+B$ 是由 l 个 9 所组成的数 $99\cdots9$。

[注 1] 本篇未指明的小数都是指十进位小数。——译注

[注 2] 不排除假因子 $\varphi(b)$ 本身。

24. 圆的一个特性

下雨则地湿，然而地湿却不一定是下雨。这是在说明一个定理和它的逆定理的区别时常用的例子。在这个规范化的叙述中，正、逆定理的区别是很清楚的，但在日常生活中却常常很容易混淆。聪明人在他们注意的时候，很明白这个区别，而在普通交往中却不自觉地会搞糊涂。政治上的雄辩家常常抓住对方的话，用相反的形式使它变得十分可笑，除非他的听众已经注意到这个诡计。每个数学家都知道，在教学时，必须向初学者系统地说明要避免一个共同性的错误，即不自觉地利用了一个未被证明的逆定理。

但是，自觉地由定理导出它的逆定理，对于数学的研究，是有用而富有成果的原理。彼此独立的本篇以及下一篇，将说明这个原理如何有助于由已知的定理引出新定理和新概念。

我们从一个简单的数学定理及其逆定理开始。圆内对同弦的所有圆周角必相等的定理[1]，是初等几何中所熟知的(图 84)。现在主要指出此定理的逆定理也是对的：在定线段 AB 上所有张角相等的角的顶点的轨迹必是圆。在初等几何中，这个定理和它的逆定理的重要性不是很明显的。由于定理和逆定理都是对的，所以圆周角就是圆的一个特性。它可以用来代替通常这样的定义，即圆是和定点等距的点的轨迹。事实上，所有有关圆的最有趣的定理，只有在引入此定理后才能证明，并且比通常的圆的定义有更直接的依赖关

系。

在谈了这些预备知识后，我们转向圆的另一性质，并证明它是圆的一个特性。在初等几何中，夹角定义为（有很好的理由认为）二条直线之间转过的角度。然而，我们没有理由不去考虑二条曲线之间转过的角度。如果我们愿意，它可以定义为二条曲线在顶点的切线之间的角（图85）。

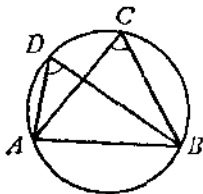


图 84



图 85



图 86

圆有一个明显的性质，即连接圆上任意二点的弦在这二点与圆的夹角是相等的。这些角是直线和曲线之间的角。我们现在要建立逆定理，并且要问：如果曲线有这样的性质，即连接它的任意二点的弦在这二点与曲线的夹角相等，那么此曲线一定是圆吗？抑或是具有此性质的另一条曲线呢？（图 86）我们将证明，此曲线一定是圆，而且这是圆的一个特性。

设有一具有此性质的曲线，A,B,C 是曲线上的任意三点（图 87）。我们作连接这些点的三条弦以及曲线在这三点的切线。于是，由假设的性质知道，图内用相同字母标出的角是相等的。

在点 A 的三个角组成一个平角，同样，在 B、在 C 点的角也是如此。所以有

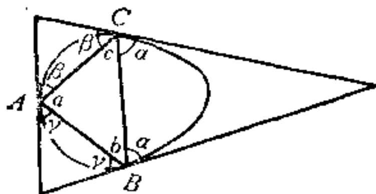


图 87

$$\angle a+\angle\beta+\angle\gamma=2R,$$
$$\angle\alpha+\angle b+\angle\gamma=2R,$$
$$\angle\alpha+\angle\beta+\angle c=2R。$$

其中R表示直角。这些等式相加，得

$$(\angle a+\angle b+\angle c)+2(\angle\alpha+\angle\beta+\angle\gamma)=6R.$$

现在，由于三角形三内角的和是$2R$，我们有

$$\angle a+\angle b+\angle c=2R,$$

所以 $$2(\angle\alpha+\angle\beta+\angle\gamma)=4R,$$
$$\angle\alpha+\angle\beta+\angle\gamma=2R,$$

与 $$\angle a+\angle\beta+\angle\gamma=2R$$比较

我们得到$\angle a=\angle\alpha$。同样的方法，得$\angle b=\angle\beta$，$\angle c=\angle\gamma$。

现在设D是曲线上的另外任意一点。我们可以作三角形ABD，并且如同三角形ABC一样来处理它。由于点A和B在二个三角形中是相同的，弦AB，过点A以及过点B的切线也就相同。由此，角γ在二个图形中是相同的。由于在D点的角$\angle d$等于$\angle\gamma$，而$\angle\gamma=\angle c$，那么，对于曲线的任意点，都有$\angle d=\angle c$。所以对于曲线上的任意点D，弦AB所张的角都是相等的。由圆周角的逆定理，这就推导出

D一定是在 A、B、C 三点所确定的圆上。由于D是曲线上任一点，所以曲线必定是圆。

[注] 如果D在AB的下侧弧上，那么角就是指线段AD的延长线和线段DB之间的角(否则必须考虑补角)。

25. 等宽度曲线

1. 圆的定义是：与一个定点，即圆心等距的所有点组成的曲线。轮子是圆的这个性质直接的实际应用。轮毂由于等长的车辐而处在地面上的一定高度上，这就得到了一个平稳的水平运动。运动中如果有很大的载重，轮和轴就不能保持十分坚固。遇到这种情形，往往依靠更原始的滚轴(图88)。重物在连续放在前面的圆棍上滚动，它在横截面为圆的棍上保持水平运动。

显然，轮子必须作成圆的形式，轮毂要在圆心处，因为其它的形式将产生一个忽上忽下的运动。然而，使人惊奇的是，对完成滚动所需要的性质来说，棍的横截面不一定非圆不可。对于滚轴来说，圆心也不再是重要的。圆用来作滚动的原因是由于圆有这样的性质，即当圆不管怎样滚动时，圆的任何一对平行切线的距离总是相等的。即圆在任意方向都有相同的宽度；因而圆也就是所谓的"等宽度曲线"。如我们上一章已讨论过的那种性质，可能猜想圆的这个性质是它的完全特性[1]。使人惊讶的是，存在着具有这种性质的其它曲线。事实上，存在着大量的非圆等宽度曲线。

2. 如果我们希望按一定方向确定某曲线C的宽度，我们可以把这条曲线的所有点垂直投影在平行于这个方向的直线上(图 89)。这些投影可以布满直线的某个线段AB，而这个直线段的长度，就是曲线在一定方向的宽度。

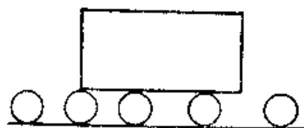


图 88

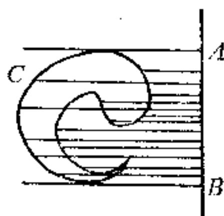


图 89

在A和B处垂直于AB的二条边缘投影线至少和曲线C有一个公共点，而整个曲线位于这种直线的一侧。我们称具有这种性质的直线是曲线的“支持线”。

一条闭曲线在每一方向，都正好有二条支持线。它们可以按图 89 的那种方法找到，或者按一定方向画二条平行线，使曲线处在它们之间，然后慢慢移动它们，直到和曲线恰好接触。(图 90)。

一条支持线和一条切线是不同的。图 91 a 中的直线t是在T处的切线，但不是支持线。图 91 b 中S是支持线，但不是切线。

对于等宽度曲线，任一对支持线之间的距离是一个固定的量b。如果我们画出曲线的二对支持线，它组成的平行四边形应是菱形(图 92)。如果二对支持线是垂直的，菱形就是边长为b的正方形。因此，所有外接于等宽度曲线的正方形都是全等的。以上这些，我们可以用下面的方法作一极好

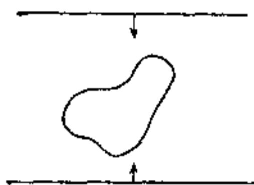


图 90

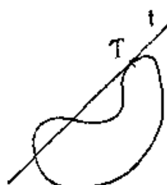


图 91a

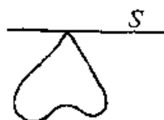


图 91b

的说明。把一硬纸卡片剪成等宽度曲线的样子，而用另一硬纸片剪下一个正方形的洞。如果正方形的边长等于曲线的宽度，那么不管方向怎样变化，它正好合适地装入曲线板。曲线图形可以在正方形内自由转动而不留任何空隙。这个命题和逆命题都是对的：等宽度曲线可以在正方形内旋转而不留任何空隙；可以在正方形内自由转动的曲线是等宽度曲线。

3．最简单的等宽度曲线不是圆，而是图 93 所示的曲边三角形。三个边是相等的圆弧，而每个圆弧的圆心是它所对的角顶。三个弧有相同的半径，这个半径等于曲线的相等宽

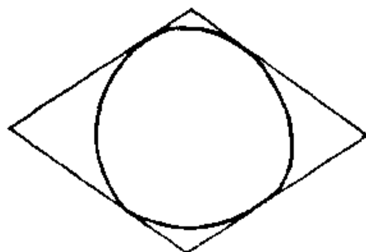


图 92

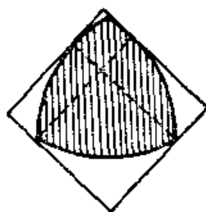


图 93

度 b。任意二条平行支持线，必有一条相交于角顶，而另一条与对边相切，或都与角顶相交。在第一种情形，二条支持线间的距离显然是 b。在第二种情形，每一条支持线都是和另一角顶所对的弧相切的，因此它们的距离还是 b。

从等宽度曲线的意义来说，这个曲边三角形是由工艺学家鲁列斯首先发现的。他从运动学的角度证明，这是一条可以在正方形内部自由转动而没空隙的曲线。我们刚才已经知道，这种性质是等宽度曲线的特性。

4. 原则上可以利用鲁列斯曲边三角形，把这个图形扩展为更多的边。关键的想法是：圆弧的心是它所对的角顶，从而画出一组有等半径的圆弧。我们开始可以把任意点 B 作为第一个角顶，以 B 为中心、b 为半径画弧。在这个弧上，我们选择 A 和 C 二点作为新角顶，以 C 为中心、以 b 为半径的弧通过 B，因为由前面的作法知 $BC=b$。在这个弧上，选择另一角顶 D。以 D 为中心、b 为半径的弧通过 C，如果我们希望结束这个过程，可以在这个弧上选择角顶 E，使它也处在以 A 为中心、b 为半径的弧上。也就是 E 是二个弧的交点。最后，用一个以 E 为中心、b 为半径的弧连接 A 和 D，这样就得到一个等宽度的曲五边形 $ADBEC$（图 94 a）。边

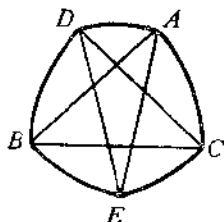


图 94a

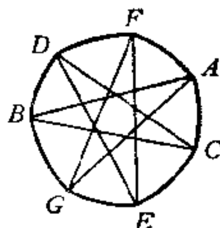


图 94b

数更多的曲边多边形，可以用同样的方法作出来，这只要多作几步，然后使曲线成为闭合的就可以了。图94 b 是这样的7边多边形。由于每一个角顶都对着一条以 b 为半径，以这个角顶为中心的圆弧，这个作法直接产生一个宽度为 b 的等宽度曲线。为了以后的目的，我们用半径把角顶和所对弧的二个端点连接起来，这样使每个角顶都连接起来了。这些半径组成一个自交多边形，而所有这些边都是相等的。每一对过一个角顶的半径组成的角，是所对圆弧的圆心角。

由这个方法作出的所有曲边多边形的边数一定是奇数的。为了弄清这一点，我们标出一个角顶和它所对的边。如果我们现在由所标出的角顶开始沿着曲线行走，我们将首先通过一个边，然后一个角顶，这样继续交错着直到我们经过一个角顶而恰好到达所标出的对边。总起来，由所标出的角顶到标出的边所经过的边和角顶的数目是相同的，设为 n。如果我们现在再由所标的角顶开始，按相反的方向沿曲线行走，我们又将经过 n 个边和 n 个棱到达所标出的边。因为第一条路径上的角顶的对边，都在第二条路径上，并且第一条路径上的边所对的角顶也在第二条路径上。再算上标出的部分，那么有 $2n+1$ 个角顶和相同数目的边。

5．我们已经作出的所有曲线都有角顶，也就是都有二个边相交为一个角的点。然而，我们可以利用这些曲线得到一种没有任何角顶的新的等宽度曲线。由一条有角顶的曲线开始，我们画一平行于它的曲线，并且以固定的距离 d 位于它的外部（图95 a, b, c）。这很容易通过在曲线内部画一个辅助的对角线多边形而作出。然后用有相同中心，半径增加了 d 的圆弧来代替原来的曲线弧。圆弧的角顶看作是半径

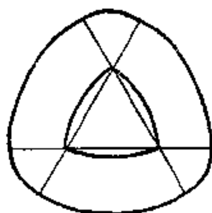


图 95 a

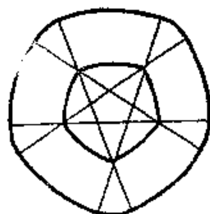


图 95 b

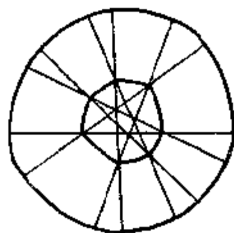


图 95 c

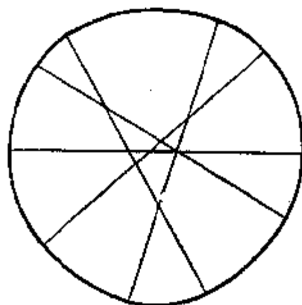


图 96

为零的弧，这样它们就由半径为 d 的弧所代替。最后的图形是由奇数个相同半径的弧，以及同样多个另一半径的弧组成的。这些弧可按如下方式配对，即一种半径的弧对应一个有相同中心(原曲线的角顶)的另一半径的弧。

同一原理可以用来构作这样的等宽度曲边多边形，这种多边形的圆弧有二种以上不同的半径。它的对边可以安排成有同一个中心，因此它们的圆心角组成对顶角(图 96)。

这些方法使我们可以构作无数个等宽度曲线。然而，这些曲线都具有特殊的形状，即它们都是由许多圆弧组成的。

为防止误解，我们要强调指出，有这样的等宽度曲线，它的一部分不管是多么小，都不是圆弧。

6．现在我们已经见过一些等宽度曲线的例子了，我们将开始研究它们的一般性质。在我们所有的例子中，曲线都是凸曲线，也就是和任意一条直线相交只可能有二个公共点的曲线。为了简化讨论，我们将限制在只研究凸曲线，甚至在我们没有明确指明是凸曲线时，说的也是这种曲线。事实上，这并不是什么限制。可以证明凡是等宽度曲线都是凸的。然而这个证明会使我们离开主题，所以我们作出这种限制而加以避免。

较精确的定义是，凸曲线是凸区域的边界。一个凸区域是按下述性质确定的，即区域内的任意二点都可以用全部在区域内的一条直线段连结起来。凸区域的例子是：正方形，圆，三角形，椭圆，以及我们已提到过的所有的等宽度曲线。显然，一个凸区域的支持线或者只有一个点，或者有一线段是和区域的边重合的。

我们将证明定理：

定理Ⅰ：等宽度曲线和它的每一条支持线只有一个公共点。

在证明这一点以前，我们来作一个简单的观察：

定理Ⅱ：**等宽度 b 曲线上任意二点的距离至多等于 b。**

如果 P 和 Q 是曲线上任意二点（图 97），那么 PQ 一定是在垂直于 PQ 的二条支持线之间。所以这二条线之间的距离，至少和 PQ 的距离一样大。由于支持线间的距离是 b，所以结论已得到证明。

至于定理Ⅰ的证明，我们假设它是错误的，即曲线上有

二个点 P_1 和 P_2 在支持线 S 上(图 98)。我们在曲线的另一侧作一条与 S 平行的支持线 S'，同时设 Q 是 S' 和曲线的交点。S 和 S' 间的距离仍是 b。

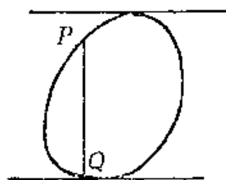


图 97

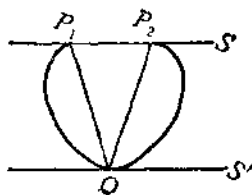


图 98

由于三角形 P_1QP_2 不能有二个直角，所以线段 P_1Q 和 P_2Q 不能同时垂直于 S。结果必有一个线段的长大于 b，但这和定理Ⅱ有矛盾。所以曲线上有二个点在支持线上的假设是不可能的，因而我们得定理Ⅰ。

如果我们再次利用这样一个事实，即连接二条支持线的垂线长度为 b，而任何其它连接线则较长，那么直接得到定理：

定理Ⅲ：如果用一条直线把等宽度曲线与二条平行支持线的二个交点连接起来，那么它必垂直于支持线。

7．如果以等宽度曲线上任一点为圆心，以 b 为半径画圆，那么由定理Ⅱ，整个曲线将被圆所包围。我们将证明曲线不可能整个都在圆的内部，它至少有一个点在圆周上。

设 P 是等宽度 b 曲线 c 上的任意一点。以 P 为圆心画一圆 K_1(图 99)，这个圆要大到能围住曲线，但又要小到 c 上有一点 Q 在它的圆周上。由于以 P 为中心，以 b 为半径的圆

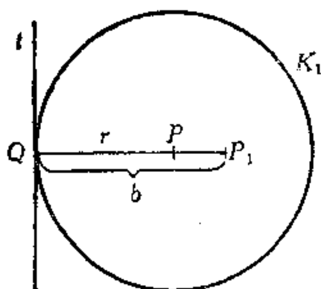


图 99

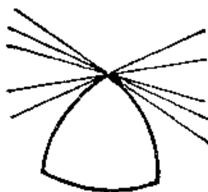


图 100

K包围了曲线c，所以K_1的半径r最大等于b，因而K_1在K的内部或最大等于K。

在Q点作圆K_1的切线t，它也通过c上的Q点。又因为c被K_1包围，所以它整个处在t的一侧。所以，t是c的支持线。在c的另一侧，有平行于t的支持线S，由于c是等宽度曲线，S和t的距离应是宽度b。根据定理Ⅲ，S和等宽度曲线的交点P_1是位于t在Q点处的垂线上的。如果$r=b$，那么P_1落在P上；如果$r<b$，那么P在Q和P_1之间。但后一种情形是不可能的。三点Q，P，P_1都属于c，而且都在一条线上。因为凸曲线被直线切割只有二个交点，所以和凸曲线有二个以上的公共点的直线，只可能是支持线。但根据定理Ⅰ，我们知道，等宽度曲线和支持线只有一个公共点。因而P_1必落在P上，那么$r=b$。

在这个证明中，P是c的任意一点，并且过此点作出了c的支持线S。因此我们已经证明了这样的结果：

定理Ⅳ：过等宽度曲线上的任一点至少有一条支持线。

一条等宽度曲线上可能有这样一些点，过这些点的支持

线不只一条。这样的点叫角顶。在以前的例子中，很多等宽度曲线都有角顶。在角顶上，很明显，在二条支持线组成的角内的直线都是支持线(图 100)。因而凸曲线在角顶有一束支持线。其中处于极端位置的二条直线是这束线的边界。

如果 P 是 c 上的任意一点，那么由定理Ⅳ，我们在这一点可以画出 c 的一束(或一条)支持线 S。过 P 画 S 的垂线。这条线交 c 于对面的点 Q，并且 PQ 有长度为 b。以 Q 为圆心、b 为半径的圆将包围 c，而且 S 将作为切线。这可以得到如下定理：

定理Ⅴ：过等宽度曲线的任意一点 P，可以作一个半径为 b 的圆，使之包围该曲线而且在 P 点与曲线支持线相切。或者与指定的一条支持线相切，如果支持线多于一条的话。

8. 下面的定理建立了等宽度曲线和圆之间的联系：

定理Ⅵ：如果圆和等宽度曲线 b 有三(或更多)个公共点，那么圆的半径长至多是 b。

鲁列斯三角形表明，这样一个圆的半径可以等于 b。如果这三个弧中任意一个扩展为整个圆，那么它和曲线有无限多个公共点。

为了证明定理Ⅵ，我们假设圆 k 和等宽度 b 曲线 c 有公共点 P,Q,R。三角形 P,Q,R 的三个角中，至少有一个角不小于其它两角，也就是它同时大于其它二角，或等于另一个角而大于第三角，或也许同时等于其它二角。我们可以假设这样的角在 P 点，并且记为 α。过 P 作等宽度曲线的（或一条)支持线。然后作半径为 b 的圆 K，使在 P 点和支持线相切并且包围C。点 Q 和 R 在 K 的内部或在 K 的圆周上。如果 Q 和 R 都在这个圆周上，那么由于过三点只有一个圆，K 和 K

是恒等的。在这种情形下，没有更多的东西需要证明。

否则，我们延长 PQ 和 PR 到它们和K的交点 Q' 和 R'（图 101）。我们现在希望证明 $Q'R'$ 大于 QR。

如果Q和Q'恰好是相同的点，那么R和R'是不同的，因为在这种情形下，Q和R是同时被固定在K上的。这里图 101 的三角形的关系如图 102^a 中所示那样。角 $QRR'=\delta$ 是三角形 PQR 的一个外角。现在，由初等几何中的定理，外角大于三角形中和它不相邻的二个角中的任意一个。这里的情形是 $\delta>\alpha$。又因为β是三角形 QRR' 的一个外角，我们有 $\beta>\beta'$。由于我们在三角形 PQR 中选取 $\alpha\geqslant\beta$，所以有 $\delta>\alpha\geqslant\beta>\beta'$，或 $\delta>\beta'$。因为 QR' 是三角形 QRR' 对角为 δ 的一边，而 QR 对的是较小的角。那么由熟知的几何定理，有 $QR'>QR$。

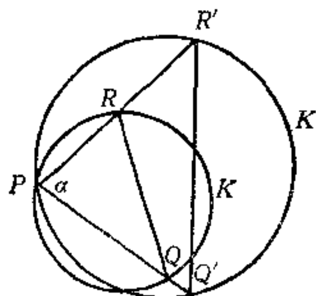


图 101

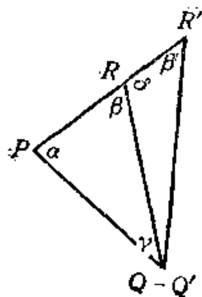


图 102a

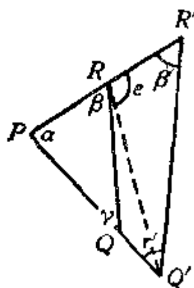


图 102b

如果Q不同于Q'而且R不同于R'，三角形的关系就如图 102 b 所示。根据三角形三内角和的定理，我们有$\beta+\gamma=180^\circ-\alpha=\beta'+\gamma'$。因而$\beta'>\beta$和$\gamma'>\gamma$同时成立是不可能的。假设第一个不等式不成立，我们有$\beta'\leqslant\beta$。在四边形$QQ'R'R$内，我们作不使β'分成两部分的对角线$Q'R$。(在$\gamma'\leqslant\gamma$的情形下，作QR'。) 用ε表示角$Q'RR'$，我们看到，它是三角形$PQ'R$的一个外角，于是有$\varepsilon>\alpha$。又因为早已有$\alpha\geqslant\beta$和$\beta\geqslant\beta'$，最后得到$\varepsilon>\beta'$。那么在三角形$Q'R'R$中，边$Q'R'$所对的角大于边$Q'R$所对的角，于是我们有$Q'R'>Q'R$。由于前已证明$Q'R>QR$，我们最后得到了所希望的结果$Q'R'>QR$。

我们已经得到当圆k和K不相同时（图 101）的各种情形，都有$Q'R'>QR$。现在α是圆k的圆周角，并且是弦QR所对的角，同时也是圆K的圆周角和弦$Q'R'$所对的角。因而QR和$Q'R'$在圆k和圆K中分别对着同样的圆心角2α。如果把这些圆心角放在一起，就得到图 103。这里我们立即知道大弦属于大圆，因而可见K的半径b大于k的半径。这就完成了定理Ⅵ的证明。

9．最简单的等宽度曲线不是圆，而是有角顶的鲁列斯三角形。下面的定理说明鲁列斯三角形由于它的角顶具有的性质，而成为等宽度曲线中比较特殊的一种。

定理Ⅶ．等宽度曲线的角顶处的角不能小于 120°。唯一的在角顶处有120°角的曲线是鲁列斯三角形，它有三个这样的角顶。

我们用在角顶的支持线束的边界支持线来度量角顶的角。如果角顶Q有角v，那么支持线束的角是$180^\circ-v$(图

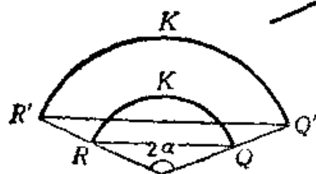


图 103

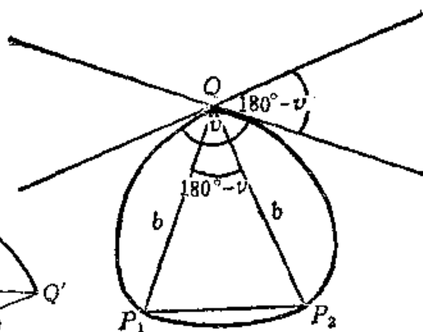


图 104

104)。在Q处这些支持线的垂线组成角度为 $180°-v$（即角 $P_1Q\ P_2$）的另外的一束线。由定理Ⅲ我们知道，每条垂线、交曲线于和Q的距离为 b 的一个点。

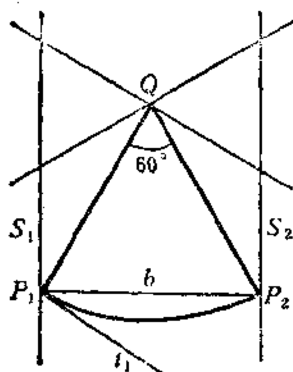


图 105

因而角顶Q所对的那部分曲线，是半径为 b 的圆弧，而且圆心角是 $180°-v$。根据定理Ⅱ，弦 P_1P_2 的长不能超过宽度 b。那么等腰三角形 QP_1P_2 有长度为 b 的腰，而底边长度不超过 b。因而角 P_1QP_2 最多为 $60°$。我们早已知道，这个角 P_1QP_2 是 $180°-\theta$，那么就有 $180°-\theta\leqslant 60°$，因而 $v\geqslant 120°$。由于Q是在任意一个角顶处的角，定理的第一部分已得到证明。

现在如果角顶的角是120°，那么角 P_1QP_2 是60°，而且等腰三角形 P_1QP_2 是等边的(图105)。那么 P_1P_2 的长度是 b。由于这个长度等于曲线的宽度，垂直于 P_1P_2 的二条支持线必定通过 P_1 和 P_2。由此我们可以知道，P_1 和 P_2 也都必定是曲线的角顶。P_1 和 P_2 之间的那部分曲线正如早已见到的，是一段圆弧。不仅 S_1 是在 P_1 处的支持线，在 P_1 处的圆的切线 t 也是支持线。在这二条线之间的夹角，很容易知道是120°。由于没有角小于120°的角顶，所以 S_1 和 t_1 必定是过 P_1 的支持线束的边界线。在 P_1(同样在 P_2)的角顶，因而正好是120°角。那么 P_1 和 P_2 就和刚才的Q有相同的性质。它们所相对的都是半径为 b 的圆弧，并且中心角是60°。但这恰好是鲁列斯三角形，因而定理Ⅶ的第二部分得到证明。

10. 在这一篇中我们首先得到了一些特殊的等宽度曲线，并且知道了怎样构作它们。然后，我们证明了由定理Ⅰ到定理Ⅶ所表示的一般性质。这些性质对于所有的凸等宽度曲线都是成立的，但是它们一点也没谈到关于等宽度曲线的存在性。我们现在将给出一个可以产生任意等宽度曲线的完全一般的构作法，这将给我们关于一切可能的等宽度曲线的完整看法。定理Ⅴ所证明的性质是特别重要的。在下述的意义下，这个性质可以说刻划了等宽度曲线的特点，即在满足定理Ⅴ的条件的曲线上，在相对的二点之间，分曲线为二段弧，我们任选一段，都可把它延拓为等宽度曲线。更确切地说，我们断定：

定理Ⅷ. 如果凸弧 Γ[8]有长度为 b 的弦，整段弧处在过弦的端点与弦垂直的二条直线之间，又如果该弧具有这样的特性，即能被以 b 为半径，与一支持线相切于Γ和支持线

的交点，并和弧同处于线的一侧的每个圆所包围，那么这段曲线可以被延拓成为等宽度 b 的曲线。

11. 在证明这个定理时，最好应用等宽度区域的概念。由于任意等宽度区域都以等宽度曲线为边界，所以我们只需证明是否存在一个相应的区域即可。

在开始论证以前，我们必须作一个关于区域的"交集"的说明。如果已给出几个区域，那么这些区域的公共部分

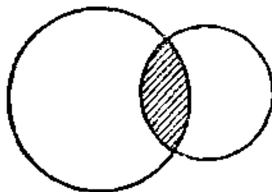


图 106

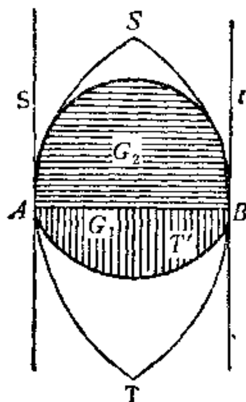


图 107

就叫作它们的交集。例如图 106 所示的二个圆的交集是阴影区域。

任意一组凸区域的交集还是凸区域。

要证明这一点，我们必须说明交集的任意二点，可以用完全在交集内的线段连接。但这是明显的。因为，如果 P 和 Q 二点是在交集内，它们就处在这些区域中的任一个中。那么，由于这些区域中的每一个是凸的，线段 PQ 就在所有的区域内。由于它在所有的区域内，线段 PQ 必然也就在交集内。

在这个证明中，这个组究竟是含有限个，还是含无限个凸区域则无关紧要。

12. 现在我们假设弧 Γ(图 107) 有定理 Ⅷ 所需要的性质；弦 AB 有长度 b。这个弧连同它的弦，是区域 G_1 的边界。过点 A 和 B，垂直于 AB 的直线是 G_1 的支持线。过支持线和 Γ 的交点，和支持线相切，并且半径为 b 的任意一圆包围了 Γ。

对于 G_1，我们附加一个以弦 AB、圆弧 AS 和 BS 为边界的新区域 ABS，其中圆弧 AS 以 B 为圆心，圆弧 BS 以 A 为圆心。这个区域我们叫 G_2。二个凸区域 G_1 和 G_2 合在一起组成凸区域 G，这个区域的边界线是 Γ 以及弧 AS 和 BS。

现在我们来研究一下所有圆心在 Γ 上，半径为 b 的圆的集合。区域 G 和这无限多个圆的集合有一个凸的交集 D (图中的阴影部分)。我们现在将证实这个区域 D 是以弧 Γ 作为它的一部分边界的等宽度区域。

如果 Γ 属于 D，它一定是在边界上，因为它原先就是 G 的边界。Γ 还在以 Γ 上的点为圆心，以 b 为半径的所有圆内。要证实这一点，需要证明 Γ 上没有二个点的距离是大于 b 的。现在由假设，Γ 同时在以 A 和 B 为圆心，以 b 为半径的二个圆内，因而它是在曲边图形 $SATBS$ 内。由于 Γ 是在 AB 的一侧，它必在作为 G_2 对于 AB 的镜面映像的区域 G_2' 内。G_2' 内任意二点的距离显然最多是 b，因此，这对于特殊的 Γ 上的任意二点也必定是正确的。因而 Γ 属于 D 而且是 D 的一部分边界。由于 D 是凸的，它必然包含连接 Γ 上的二点的所有弦。由此 G_1 必是 D 的一部分。

D 中没有二个点的距离能够大于 b。由于 D 是 G 的一部

分，而G是由G_1和G_2二个区域组成，我们有三种情形需要研究：如果二点都在G_1内，那么如前所述，它们的距离不能大于b。如果二点都在G_2内，那同样是对的。最后，如果有一点P_1在G_1内，另一点P_2在G_2内，我们可以连接P_1和P_2，并且延长直线直到和Γ相交于P点。这三点按照PP_1P_2的次序同在这条直线上。以P为圆心，以b为半径的圆包含整个D，因此它包含这三个点。所以P_1和P_2在半径为b的圆内，因而它们之间的距离不能超过b。

我们刚才所得的结果表明，区域D的不论哪个方向的宽度都不能超过b。我们现在必须证明D的任意方向的宽度都是b。AB方向的宽度，定理已经指明是b。现在研究任一其它的方向，并作出垂直于这个方向的D的二条支持线。二条中的一条，比如S_1，将有一个与Γ的公共点Q。在Q处作一条长度为b的S_1的垂线，并且把它的终点叫作M。而M属于D。要证明这一点，我们必须说明，M是在G内，同时也是在以Γ上的点为中心，以b为半径的所有圆内。这后一点需要我们证明M和Γ上的每一点距离最多是b。这可以由以M为中心，以b为半径的圆和S_1切于Q的事实得到。根据假设，这个圆必包含弧Γ，而这就证明了由M到Γ的任意点的距离不大于b。在特殊情形下，距离AM和BM不大于b，所以点M必在图形$SATBS$内。又因为M是在与Γ相对的AB的另一侧，所以它也在G_2内。因而它是在G内，同时也是在所有的圆内。所以最后，M在交集D内。

由于QM垂直于S_1和S_2，而Q和M属于D，所以支持线S_1和S_2之间的距离，必然至少是和$QM=b$一样大。这个距离不可能大于b，因为这会意味着二个交点之间的距离大

于 b，而我们已知，这对于D内任意二点是不可能的。因此S_1和S_2之间的距离恰好是 b；区域D的任意方向的宽度都是 b。

这个定理证明：满足一定要求的弧 Γ 是可以延拓成为等宽度 b 的曲线C的。可以看出，只有一种方法可以使 Γ 扩展组成为等宽度曲线，因而C是唯一确定的。

13．最后，我们不加证明地指出这些曲线的另一个值得注意的性质：有相同宽度 b 的所有等宽度曲线，有相同的周长。而且其周长必等于直径为 b 的圆周长。这个事实极易用第 4、5 节所作的例子验证，而这些例子都是利用有相同圆心角的类似的圆弧作出的。然而，一般曲线的证明，需要用到超出本书范围的概念和方法。这个证明只有在极为仔细的分析了曲线长度的概念后才能开始。

[注 1] 即这个性质是圆的充分必要条件——译注

[注 2] 即一段弧，它和它的弦一起，是凸区域的边界。

26．初等几何作图中圆规的不必要性

1．初等几何中的作图，都是用直尺和圆规完成的；甚至按这个原则把初等几何从整个几何学中区分出来，即认为，初等几何是能用直尺和圆规作图的那部分。但是这两个工具并不是完全必需的。许多作图题只需要用这一个或那一个工具就可以解决。而且，按照马歇罗尼的研究以及最近发现的

莫尔的早期著作所说，指出直尺是完全不必要的。所有能用直尺和圆规完成的作图，可以完全不要直尺而仅用一个圆规来完成。由于没有直尺，就画不出直线，在这种情形下，认为只要找到二点，一根直线就已经作出来了。斯坦纳从另一方面发现，只要预先作出某个固定的圆以及它的圆心，初等几何中所有作图题仅用直尺就能完成。证明这个定圆是必需的这一点并不困难。我们将证明，对于没有给出圆心的固定圆，只用直尺来完成所有作图题是不可能的。实际上，未给出圆心的二个不相交的圆，仅用直尺完成所有作图是不可能的。但是，我们已经知道，没有给出圆心的二个相交的圆，或三个不相交的圆，就完全可以代替斯坦纳的给出圆心的圆了。

2．由于根据斯坦纳的结果，只要我们给出了一个圆心已知的圆，所有初等几何作图题都可以作出，这就需要证明，不论是已知一个没有给出圆心的圆，还是二个不相交的圆，仅用直尺要找到圆心是不可能的。也就是说，要证明这二个作图题是不可能的。我们的证明是，首先反设仅用直尺可以作出这个圆的圆心或两个不相交圆的圆心，然后引出一个荒谬的结果，从而证明了这种作图是不可能的。这些间接证法将利用映射的原理。

如果我们证明了只用直尺不可能求二个不相交圆的圆心，那么很明显，也就不可能在已知一圆的情况下求得圆心。但后一种情形在几何学方面是比较简单的，因此我们将首先证明它。而且它对于介绍这二个证明所应用的思想还有所帮助。

3．假设对于已经画在纸上的圆，我们会用某种办法仅

利用一个直尺求出圆心。我们的作图，不外是作出与圆周相交的或彼此相交的一些线，以及把已经求得的交点连接起来的一些线。于是，作出的整个图形，应当由已知圆和若干直线组成，这些直线中的二条交于圆心。

现在我们来研究对于这整个图形的一类特殊的映射。通过这个映射，使假定的圆变成一个圆，每条直线变为直线，二条直线的交点变为映射下相应的二条直线的交点。这类映射显然是很多的。例如，图形成比例地放大或缩小就是这样一种映射。但是，这种相似的映射，并不适合我们的要求。我们要找的映射是这样一种映射，就是它把圆变成圆，一切直线变成直线，不过要使图形发生畸变。特别是要使圆心映射为不是象圆圆心的一个点。

只要我们找到了所希望的这种映射，我们的证明实际上就完成了。映射后的图形与原来的图形无论怎样的不同，就所涉及的作图来说，这二个图形完全是相同的。原证明的每一步，作一直线，找一个交点，连结两个交点，在像图上都能一步一步地完成。但是，原有圆的中心的映像，并**不是**映射圆的中心。因此，原图上与圆心相交的直线映射后的像直线，不交于像圆的圆心上。虽然整个作图在映射后的图形上一步步都能完成，但却不能找到像圆的圆心。这与能确定圆心的作图假设有矛盾。因此，这就证明了仅利用直尺求作已知圆的圆心是不可能的。

至于二个圆的情形，证明是十分相似的。

4．现在我们必须找到我们所叙述过的那种映射。我们将借助空间的射影得到这样的映射。取图形(图 108) 所在平面 E 以外的一点 O，并画出另一平面 E' 作象平面或投影平

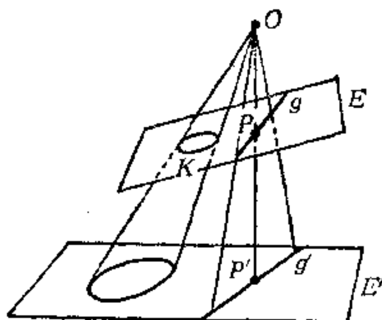


图 108

面。从点 O 出发过平面 E 的某个点 P 引射线，它必然交平面 E' 于点 P'。于是 P' 就是点 P 的像或投影。同样的方法，E 上的整个图形，能逐点的投射为平面 E' 上的图形。这个投影可设想为由点光源 O，把平面 E 上的图形投射到 E' 平面上。通过投影，任意直线 g，将映射为直线 g'。事实上，所有过 O 及 g 上的点的射线的集合，组成一个平面。这个平面与平面 E' 的交是直线 g'。

圆的投影一般并不都是圆。过 O 及圆 K 上的点的所有射线的集合，组成一个圆锥。在一般情况下，这是一个斜圆锥。连接顶点和底圆圆心的直线垂直于底面的圆锥，称为"直"圆锥，其它的称为斜圆锥。我们知道，投影面 E' 和圆锥相交得到的二次曲线（标准截线），一般说不是圆。但对于我们的目的来说，圆必须再映射为圆。有二种特殊安排的映射是能作到这一点的。

第一种情形是很普通的。方法是把平面 E 与 E' 彼此平

行放置。通过投影得到的这个映射，显然是相似映射，视平面 E' 距 O 较平面 E 远或近而为相似放大或相似缩小。这样作图对我们的目的来说是不适合的，因为它不能使图形发生畸变。

第二种情形与一个立体几何的定理有关。为了不使这里的推理过程中断，我们且把对它的证明放到本篇的最后一段(第 8 节)。过顶点 O 作一平面，使它垂直于斜圆锥的底面，且过底圆圆心。这个平面是斜圆锥的对称平面。图 109 表示圆锥和这个平面的截面。只有底圆的直径 K_1K_2 表示在图上。底平面是垂直于图形的平面的。由于图形平面的位置，OK_1 是过 O 和底圆圆周上的点的线段中最短的一条，而 OK_2

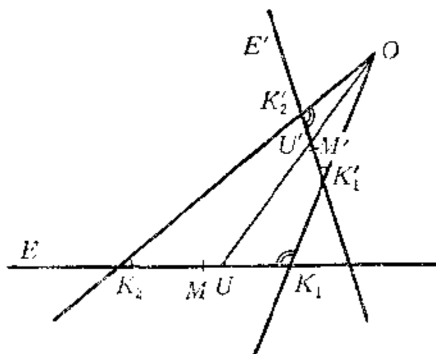


图 109

是最长的一条。每一个平行于底平面 E 的平面，显然截圆锥为一圆。如果作一平面 E'，使它与直线 OK_1 和 OK_2 交于点 K_1' 和 K_2'，且使 $\angle OK_1'K_2'=\angle OK_2K_1$，那么由第 8 节证明的定理，就能断定 E' 截圆锥为一圆。由于 $\angle OK_1'K_2'=$

$\angle OK_2K_1$，且由于三角形三内角和等于180°，所以$\angle OK'_2K'_1=\angle OK_1K_2$。显然，任意与平面$E'$平行的平面，截圆锥后的截线也是圆。

如果平面E'已按所希望的选择好，那么由E到E'的投影，就具有我们所需要的一切特点。为此，我们只要证明K_1K_2的中点M并不投影在$K'_1K'_2$的中点M'就可以了。首先，三角形K_1OK_2和$K'_1OK'_2$在顶点O有公共的角平分线。因为在任一三角形中，任一个角的角平分线，都把对边分为与邻边成比例的线段，而在斜圆锥内，我们有$OK_2>OK_1$，所以$K_2U>UK_1$，其中，U是E和角O的角平分线的交点。而且，由于三角形中大边对大角，知$\angle OK_1K_2>\angle OK_2K_1$。从而$\angle OK'_2K'_1>\angle OK'_1K'_2$，因此，$OK'_1>OK'_2$，由此我们得到$U'K'_1>K'_2U'$。由于$M$是$K_2K_1$的中点，$M'$是$K'_2K'_1$的中点，可知$M$和$M'$一定分别位于角$O$的角平分线的两边。因此，在所作的投影下，$M'$不可能是$M$的像。

5. 综合以上所述，我们已经证明了只用直尺是不能找到给定圆的圆心的。利用图109，我们可以扼要地叙述证明的要点。平面E上由圆K_1K_2和某些直线组成的图形，由O投影到平面E'上。在此投影下，直线变成直线，圆K_1K_2变成$K'_1K'_2$，但K_1K_2的中心M并没有变成$K'_1K'_2$的中心M'。也就是在平面E内任何建立在直线基础上的，并且确定了已知圆的圆心的作图，映射到E'内后将不再是这种能确定圆心的作图。所以这样考虑的作图是不可能的。

6. 在已知二个圆的情形下，这样的映射较难获得。由点O出发的投影组成两个斜圆锥，我们必须使投影平面E'截割它们时都得到圆。

我们分两种情形来讨论。首先，一个圆可能处在另一圆的内部。我们在垂直于E，并过两个圆的中心M和N的平面上，作一图形（图 110）。两个圆可由图形上它们的直径K_1K_2和L_1L_2表示。现在如果我们能找到这样的点O，使得$\angle K_1OK_2$和$\angle L_1OL_2$有同一条角平分线，那么我们可以按$\angle L_2WO=\angle L_1'W'O$的办法，且利用图 110 的符号，作出平面$E'$。因此图形内所有用同样的数字或字母标出的角都是相等的。所以根据第 8 节的定理，平面E'同时截割两个以圆$K_1'K_2'$和$L_1'L_2'$为底的圆锥。

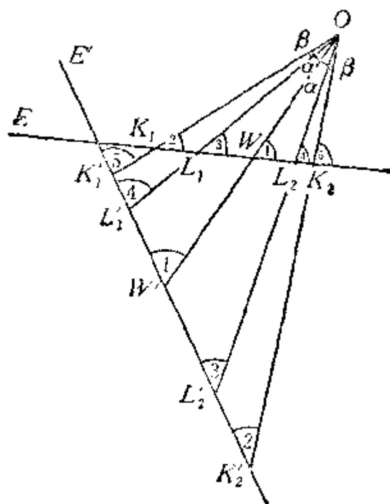


图 110

证明的其余部分和前面所述的相同：这种投影使二个圆变成两个圆，并使直线成为直线，但却不能使原圆的圆心变成象圆的圆心。在平面E内，已确定圆心的任意直线的作图

在平面 E' 上将不再如此，并且一般来说这样的作图是不可能的。

我们还必须说明，使$\angle K_1OK_2$和$\angle L_1OL_2$有相同的角平分线的点O是能找到的。假定已按此选择好O，则必有$\angle K_1OW=\angle K_2OW$以及$\angle L_2OW=\angle L_1OW$。它们相加，就有$\angle K_1OL_2=\angle K_2OL_1$。设这些角的值为$\delta$，并作底角为$90°-\delta$的等腰三角形$K_1C_1L_2$和$L_1C_2K_2$(图 111)。以$C_1$为心，过$K_1$和$L_2$作圆弧。以$C_2$为心，过$L_1$和$K_2$也作圆

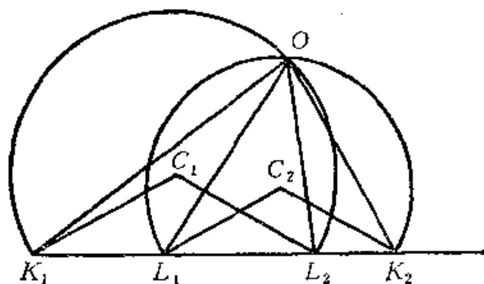


图 111

弧。由于它们的弦K_1L_2和L_1K_2有一部分重合，所以这二个圆弧必相交。设O是它们的交点。现在有

$$\angle K_1OL_2=\frac{1}{2}\angle K_1C_1L_2=\frac{1}{2}[180°-2(90°-\delta)]=\delta$$

$$\angle L_1OK_2=\frac{1}{2}\angle L_1C_2K_2=\frac{1}{2}[180°-2(90°-\delta)]=\delta。$$

因而$\angle K_1OL_2=\angle L_1OK_2$，由此，求得$\angle K_1OL_1=\angle K_2OL_2$。所以，$\angle L_1OL_2$的角平分线同样是$\angle K_1OK_2$的角平分线，这正是我们所需要的。

7．两个圆的第二种情形是，已知的两圆是彼此外离的。这里两个圆锥也是彼此分离的，并且对于任意位置O，使角K_1OK_2和L_1OL_2有相同的角平分线是不可能的。

在这种情形下，我们要利用对顶圆锥（图112）。如果$\angle L_1VO=\angle L_2'V'O$，并且$\angle K_2UO=\angle K_1'UO$，第8节的定理就适用于这两个圆锥。那么三角形$UVV'$将是等腰三角形，$UO$将平分它的顶角。所以$UO$垂直于$VV'$，并且我们有

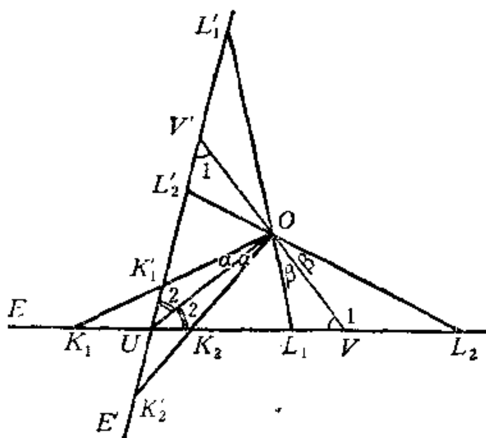


图 112

$$\angle K_1OL_1=90^\circ+\alpha-\beta,$$

$$\angle K_2OL_2=90^\circ-\alpha+\beta,$$

由此

$$\angle K_1OL_1+\angle K_2OL_2=180^\circ. \tag{1}$$

所以点O的选择必须使（1）式能够成立。也就是如果（1）成

立，那么角 K_1OK_2 和角 L_1OL_2 的角平分线，实际上是互相垂直的，这是因为

$$\begin{aligned}2\angle UOV &= 2\angle UOK_2 + 2\angle K_2OL_1 + 2\angle L_1OV \\ &= \angle K_1OU + \angle UOK_2 + \angle K_2OL_1 + \angle K_2OL_1 \\ &\quad + \angle L_1OV + \angle VOL_2 \\ &= \angle K_1OL_1 + \angle K_2OL_2 = 180^\circ,\end{aligned}$$

所以 $$\angle UOV = 90^\circ。$$

如果如图所示，在这个情况下，我们可以把 E' 放置在这样一个位置上，即用 2 标出的角是相等的，那么用 1 标出的角也是相等的。所以 E' 截两个圆锥为圆，这个投影具有所需要的性质，至于其余部分的证明和前面所述完全一样。

为了求能满足(1)的点 O 的位置，我们选择任意值 $\angle K_1OL_1 = \varphi$ 和 $\angle K_2OL_2 = \psi$，并限制 $\varphi + \psi = 180^\circ$。于是由第 6 节末尾所用的方法，我们可作一圆弧，使弦 K_1L_1 所对的内接于圆弧的角是 φ。类此，作另一圆弧使弦 K_2L_2 所对的圆周角为 ψ。由于 K_1L_1 和 K_2L_2 有一部分重合，那么圆弧将相交。它们的交点 O 就是我们所需要的。

我们现已证明：两个未给出圆心的不相交的圆，对于可以仅用直尺作出的所有基本作图来说，是不够的。

8．现在还剩下由第 4 节推延下来的那个定理的证明。这个定理断定，**如果平面** E (图 109) **截割斜圆锥** K_1OK_2 **为一个圆，那么，如果角** $K_1'K_2'O$ **和** K_2K_1O **相等，平面** E' **也将截割圆锥为一个圆。**必须记住，过顶点 O 和底圆圆心的图形的平面是和圆锥的底面垂直的。条件 $\angle K_1'K_2'O = \angle K_2K_1O$ 显然可用 $\angle K_2'K_1'O = \angle K_1K_2O$ 或 $\angle K_2'U'O = \angle K_1UO$ 代替。

图 109 的平面是圆锥的对称平面。我们的证明决定于寻求另一个对称平面。

图 113 表示的圆锥的截线与图 109 相同。

图 113

图 114

我们画一个三角形 OK_1K_2 的外接圆以及角 O 的角平分线 MO。由于角 K_1OM 和角 K_2OM 相等，K_1M 和 K_2M 的短弧也相等。所以过圆心 C 和弦 K_1K_2 垂直的直线交圆于点 M。如果把圆围绕轴 $M'CM$ 旋转，就描画出一个球。圆锥的底圆的圆周 K_1K_2，及顶点 O 在此球面上，即圆锥内接于此球。所以点 M 与圆周 K_1K_2 上的所有点是等距的。

直线 OM 在斜圆锥的研究中起着特殊的作用。它被叫作**圆锥的轴**。如果一个平面通过此轴，它就与圆锥相截得一顶点在 O 的三角形，而其它二点 H_1 和 H_2 则在圆锥的底面上。此平面截球得一圆，该圆外接于三角形 OH_1H_2。点 M 也在此圆上，并且我们得到的图形(图 114) 完全和图 113 类

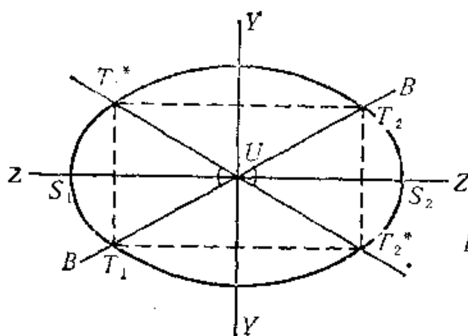


图 115

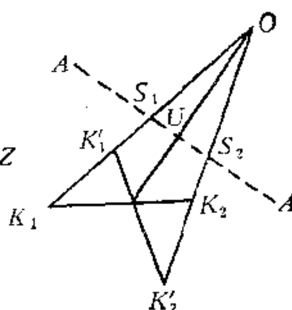


图 116

似。由于M和圆锥底圆的关系，二弦 H_1M 和 H_2M 相等，因此轴 OM 是角 H_1OH_2 的角平分线。斜圆锥的轴有这样的性质，即过此轴的每个平面截圆锥为一三角形，而在此三角形内，该轴就是一个角的角平分线。

现在我们用垂直于圆锥的轴和图 113 的平面的平面A截割此圆锥。平面A和图形平面的交线如图 113 的点线 AA 所示。A和圆锥的交线表示在图 115 上。这个图形所在的平面是平面A，并且圆锥轴垂直于它。图 113 的平面截图 115 的平面为直线ZZ。我们假设通过轴的一个平面，它截图 115 的平面于某条线，设为 BB。点 T_1 和 T_2 表示它和曲线的交点。顶点O不在图 115 的平面上，但它位于点U的正上方。在三角形 T_1OT_2 内，轴 UO 垂直于 T_1T_2，并且由我们上面的结果，它是角 T_1OT_2 的角平分线。所以我们有 $T_1U=T_2U$。由于 B 是过轴的任一平面，这就意味着点U是圆锥和平面A的截线的对称中心。

图 113 的平面是圆锥的一个对称平面，并且它和图 115

的平面的截线是 ZZ。所以图 115 对于线 ZZ 和点U 都是对称的。换句话说，圆锥上的 T_1^* 和 T_2^* 是对应于圆锥上的 T_1 和 T_2 的镜像。

这四个点组成了四边形 $T_1T_2^*T_2T_1^*$。由于 T_1U 和 T_2U 是相等的，则其镜像 $T_1^*U=T_2^*U$。而且，$\angle T_1UT_1^*=\angle T_2UT_2^*$，因为它们是对顶角。由此，三角形 $T_1^*T_1U$ 和 $T_2^*T_2U$ 是全等的，所以 $T_1T_1^*=T_2T_2^*$。因为直线 $T_1T_1^*$ 和 $T_2T_2^*$ 同时垂直于 ZZ，所以它们也是平行的。这样，$T_1T_2^*T_2T_1^*$ 是一个平行四边形。又因为二条对角线 T_1T_2 和 $T_1^*T_2^*$ 相等，所以这个平行四边形是矩形，U 则是它的中心。这表明图 115 也有对称轴 YY。

我们的结果表明，过圆锥轴且垂直于图 113 和图 115 所在平面的平面，是圆锥的第二个对称面。需要记清的是，该平面包含了图 115 中所示的直线 YY，而此直线过U且垂直于图 113 的平面。

如果我们令圆锥对第二对称平面作反射，它的像仍是原圆锥；这恰好是叙述对称性的另一种方法。这样，圆 K_1K_2 的像是另一个圆 $K_1'K_2'$，它必然也在此圆锥上(图 116)。这正是我们所期望的圆。这个圆所在的平面和圆锥轴之间的夹角和定理的约定是一致的。最后，我们注意到圆 K_1K_2 不能和它的像重合，因为这只有当它所在的平面垂直于圆锥的轴时才能发生，但这和圆锥是斜的假设矛盾。

27. 数 30 的一个性质

10 或 21 都不是素数。但是在 $10=2\cdot5$ 和 $21=3\cdot7$ 中，没有一个因子是同含于两数的。因此，它们被叫作“互素”的数。数 6 和 10 因为有公因子 2，所以它们不是互素的。

所有由 1 到 9 的数中，与 10 互素的是 3,7,9。9 与 10 虽然互素，但它不是素数。数 12 与 10 不同。由 1 到 11 的数中，与 12 互素的数只有 5,7,11，但它们都是素数。可以很容易地看出，数

3,4,6,8,12,18,24,30

都有 12 所具有的性质。

在具有比它小，且与它互素的所有数都是素数性质的数中，30 是最大的吗？本篇将证明正是如此。

我们首先来看如何着手来寻找具有这种性质的数。由 4 开始，往后每一个这样的数 N 必能被 2 整除。因为如果它是奇数，则它必与 4 互素，而 4 却不是素数。用同样的方法，由 9 开始往后任意一个这样的数必能被 3 整除，由于它已经能被 2 整除，它就必能被 $2\cdot3=6$ 整除。继续这样的论断，得到如下的表：

由 4 开始往后的数 N 必能被 $2=2$ 整除
由 9 开始往后的数 N 必能被 $2\cdot3=6$ 整除
由 25 开始往后的数 N 必能被 $2\cdot3\cdot5=30$ 整除
由 49 开始往后的数 N 必能被 $2\cdot3\cdot5\cdot7=210$ 整除
由 121 开始往后的数 N 必能被 $2\cdot3\cdot5\cdot7\cdot11=2310$ 整除。

在 4 与 9 之间，N的值只可能是 4,6,8；在 9 与 25 之间，仅 12，18，24；在 25 与 49 之间，只有 30(能被 30 整除的最接近 30 的数是 60，已大过 49)在 49 与 121 之间，已不可能有这样的数，因为 210 已经大于121。现在我们看到，如用同样方式使此表继续列下去，那么，13^2 是小于 $2\cdot3\cdot5\cdot7\cdot11$ 的，17^2 是小于 $2\cdot3\cdot5\cdot7\cdot11\cdot13$ 的，以此类推，于是 30 就是具有所需性质的最大的数。如第一篇那样，如果我们用符号 $P_1, P_2, P_3, P_4, P_5, P_6, P_7\cdots$ 表示逐渐增长的素数 2,3,5,7,11,13,17,$\cdots$，那么我们必须证明

$$P_{n+1}^2 < P_1\cdot P_2\cdot P_3\cdots P_n \qquad (1)$$

对于由 4 开始的所有的 n 都是对的。

在第一篇里一再提到的欧几里德的证明给出

$$P_{n+1} < P_1\cdot P_2\cdots P_n,$$

我们现在所需要证明的是

$$P_{n+1} < \sqrt{P_1\cdot P_2\cdots P_n},$$

这个论断比欧几里德的结果要强。不等式(1)也许比原来的有关 30 的问题更有意思。有了不等式(1)，原来的问题也就立即解决。因此，我们现在集中来研究式(1)的证明。

实际上小于(1)式右端的素数，比不等式(1)所断定的要多得多：不只是 $P_5=11$ 的下一个素数，即 13 小于 $\sqrt{2\cdot3\cdot5\cdot7\cdot11}=\sqrt{2310}=48.06\cdots$，17,19,23,$\cdots$等也都小于这个数。当我们对(1)继续考察下去时，这个矛盾变得更大。但是，由于素数序列极端不规则，不等式(1)的结果推广为对一切素数都成立是极困难的。契比雪夫利用极为广泛的方法，证明了 P_n 的下一个素数小于 $2P_n$，即 $P_{n+1}<2P_n$。这当然比关系式(1)要强。于是产生了这样一个问题，即是

否较弱的不等式(1)不能用初等方法解决。

邦塞在求学时期发现了(1)式的一个极巧妙的证法。它不仅不需要契比雪夫用到的一切分析方法和无限过程，而且只需应用最简单的数学观念。

1. 证明的基本思想类似于第一篇给出的欧几里德的证明。我们把前 n 个素数组成的式子

$$N=P_1P_2\cdots P_n+1 \quad 或 \quad M=P_1P_2\cdots P_n-1$$

换成只用前 i 个素数 $P_1, P_2, \cdots, P_i$，并且组成 P_i 个式子

$$M_1=P_1P_2\cdots P_{i-1}\cdot 1-1,$$
$$M_2=P_1P_2\cdots P_{i-1}\cdot 2-1,$$
$$M_3=P_1P_2\cdots P_{i-1}\cdot 3-1,$$
$$M_4=P_1P_2\cdots P_{i-1}\cdot 4-1,$$
$$\cdots\cdots$$
$$M_{P_i}=P_1P_2\cdots P_{i-1}\cdot P_i-1.$$

就象欧几里德表示式的情形那样，我们可以断定：

(a) 不论这些式子 $M_1, \cdots, M_{P_i}$ 中的哪一个，都不能被 $P_1, P_2, \cdots, P_{i-1}$ 中的任一个素数整除。

(b) 这些式子中至多有一个能被 P_i 整除。因为如果这些式子中的某二个，例如 $P_1\cdots P_{i-1}x-1$ 和 $P_1\cdots P_{i-1}y-1$ 能被 P_i 整除，那么它们的差 $P_1P_2\cdots P_{i-1}(x-y)$ 也能被 P_i 整除。由于 P_i 不能整除前 $i-1$ 个因子中的任何一个，因而应能整除差$(x-y)$。但 x 和 y 是在数列 $1,2,3,4,\cdots,P_i$ 之中，因此这样二个数的差，甚至当取它们中最小的 1 和最大的 P_i 时，也仅等于 P_{i-1}，所以，一切这样的差都小于 P_i。这样，由于较小的数不能有较大的数作其因子，所以 P_i 不能整除这些差。用同样的论证，我们可以断定，

(c) 这些式子中至多有一个能被 P_{i+1} 整除，至多有一个能被 P_{i+2} 整除，…，至多有一个能被 P_n 整除。

现在如果数 $P_i, P_{i+1}, \cdots, P_n$ 的个数小于式 $M_1, M_2, \cdots, M_{P_i}$ 的个数，换句话说，如果我们有

$$n-i+1<P_i, \tag{2}$$

那么，式M中至少有一个不能被数 $P_i, P_{i+1}, \cdots, P_n$ 中的任何一个整除。证明中的这个重要一步，可由(b)和(c)直接得到。如果我们称这种特殊的式子为 M_k，那么由命题(a)，此式不能被数 $P_1, \cdots, P_{i-1}$ 中的任一个整除，所以，它不能被前 n 个素数中的任一个整除。

下一步就是欧几里德的证法。式 M_k 可能恰好是素数，也可能被分解为素因子。即存在一个素数 P，它或等于 M_k，或可整除 M_k。由于 M_k 不能被 $P_1, P_2, \cdots, P_n$ 中任一个整除，那么这个素数一定超过 P_n。P_n 的下一个素数是 P_{n+1}，所以有 $P_{n+1}\leqslant P$。又因为 P 可以整除或等于 M_k，我们有 $P\leqslant M_k$。因为所有的M式中最大的是 M_{P_i}，那么，把这些不等式结合在一起，得到

$$P_{n+1}\leqslant P_1\cdots P_{i-1}P_i-1<P_1P_2\cdots P_i.$$

这样，上面的结果可简述如下：如果关系式(2)成立，那么我们有

$$P_{n+1}<P_1P_2\cdots P_i. \tag{3}$$

2. 显然，第1节的结果是欧几里德证法的结果

$$P_{n+1}<P_1\cdots P_n$$

的改进，因为数 i 小于 n，右边减小了。究竟 i 是多少呢？条件(2)不允许 i 为任意小的值。我们必须让 i 足够大，这样 $P_i, \cdots, P_n$ 的个数，即 $n-i+1$，就会小于这些数的第一

个，即 P_i。这个要求由于 i 本身要来回试取而显得复杂了。下面用一个简单的例子来帮助理解这个条件是怎么回事。

设 $n=5$，这样我们来考虑前 5 个素数 2,3,5,7,11。如果选取 $P_i=3$，那么若 $i=2$，则数组 $P_i, \cdots, P_n$ 包含 3,5,7,11。这个数组包含的个数 $n-i+1=5-2+1=4$，不小于它们中的第一个(4 不小于 $P_i=3$)，所以 i 选取的值太小了。如果使 i 值增加 1，即 $i=3$，则 $P_i=5$，数组 $P_i, \cdots, P_n$ 只有 3 个数，5,7,11。由于 3 小于 5，所以 i 选取的这个值是合适的。显然，i 选取任一更大的值也满足(2)。

我们现在要证明的命题是：如果 i 取满足条件(2)的尽可能小的值，那么我们将有：

$$P_1 \cdots P_i < P_{i+1} \cdots P_n。 \qquad (4)$$

当 $n=5$ 时，这个命题是正确的，因为由乘法可以验证 $2 \cdot 3 \cdot 5 < 7 \cdot 11$。为了说明当 n 增加时它也是正确的，我们必须看清，当 n 增加时最小的 i 如何变化。

当 $n=5$，我们有 $i=3$，$P_i=5$。数组 $P_i, \cdots, P_n$ 包含 3 个数：5，7，11。如果由 $n=5$ 改为 $n=6$，那就增加了一个素数 13。然而，不需要改变 i，因为数组 $P_i, \cdots, P_n$ 可以包含 4 个数。当我们改为 $n=7$ 时，那就加入了素数 17。我们现在必须增加 i，因为否则数组 $P_i, \cdots, P_n$ 将会包含 5 个数，而 5 不小于 $P_i=5$。因此如果 $n=7$，必须选取 $i=4$，那么 $P_i=7$，这就允许有 6 个数 7,11,13,17,19,23 的数组。从而对于 $n=7,8,9$，$i=4$ 是足够的。当 $n=10$ 时，i 的值必须增加到 $i=5$。那么 P_i 将是 11。因为这次 7 以后最近的素数不是比它大 2 而是大 4，所以这个 i 值对后面的 5 步，即 $n=10,11,12,13,14$ 都是足够的。下表显示出了 i 的值。(P_i 用

粗体字表示)

$n=5$：2,3,**5**,7,11

$n=6$：2,3,**5**,7,11,13

$n=7$：2,3,5,**7**,11,13,17

$n=8$：2,3,5,**7**,11,13,17,19

$n=9$：2,3,5,**7**,11,13,17,19,23

$n=10$：2,3,5,7,**11**,13,17,19,23,29

$n=11$：2,3,5,7,**11**,13,17,19,23,29,31

$n=12$：2,3,5,7,**11**,13,17,19,23,29,31,37

$n=13$：2,3,5,7,**11**,13,17,19,23,29,31,37,41

$n=14$：2,3,5,7,**11**,13,17,19,23,29,31,37,41,43

$n=15$：2,3,5,7,11,**13**,17,19,23,29,31,37,41,43,47

…………………

每次当 i 增加 1 时，素数 P_i 至少增加 2，这样 i 在 n 以后三个值中仍然相同。如果 P_i 增加的值大于 2，那么 i 的值对于更多的 n 的值保持不变。

现在我们已经看清当 n 增加时，最小的 i 是怎样变化的，所以我们就可以讨论(4)了。我们早就知道，当 $n=5$ 时(4)式正确，

$$2\cdot 3\cdot 5<7\cdot 11, \tag{5}$$

当 $n=6$ 时，i 不改变，那么(4)式就是

$$2\cdot 3\cdot 5<7\cdot 11\cdot 13, \tag{6}$$

显然这是正确的，因为(5)的右端增加了。

当由 $n=6$ 变为 $n=7$ 时，由于 i 的值改变，事情就不那么简单了。数 7 由右边移到左边，而在右边出现了新的素数 17。我们希望证明的是

$$2\cdot3\cdot5\cdot7<11\cdot13\cdot17。\tag{7}$$

不通过实际的计算，我们不能由(6)直接得到(7)。我们可以在(6)式左端乘以因子 7，而在右端乘以更大的因子 17，但这样会得到

$$2\cdot3\cdot5\cdot7<7\cdot11\cdot13\cdot17,$$

其中右边出现了不需要的因子 7。

如果我们由(5)开始，那么不经计算就能验证(7)。为此，在不等式(5)左端乘以 $7\cdot7$，而右端乘以 $13\cdot17$；得到

$$2\cdot3\cdot5\cdot7\cdot7<7\cdot11\cdot13\cdot17,$$

如果消去 7，我们有不等式(7)。

这个由 $n=6$ 到 $n=7$ 的论证没有任何实际的计算，它仅根据倒退二步到 $n=5$ 这一事实。类似的由 n 的一个值到下一个值的每一步，都可以实行完全一样的论证。它只要求当 i 值增加时，至少对于 n 的依次二个值 i 仍是相同的（我们已经看到，至少 n 的依次三个值，甚至更多的值，i 是相同的）。因此，不等式(4)对于一切 $n=5,6,\cdots$ 都是成立的。

3．不等式(4)二端乘以 $P_1\cdot P_2\cdots P_i$，得到

$$(P_1\cdot P_2\cdots P_i)^2<P_1\cdot P_2\cdots P_n \text{ 或}$$

$$P_1\cdot P_2\cdots P_i<\sqrt{P_1\cdot P_2\cdots P_n}。\tag{8}$$

这个结果再结合不等式(3)，就得到当 $n=5,6,\cdots$ 时的关系式(1)，这就是我们要证明的。而当 $n=4$ 时，通过乘法很易验证(1)是对的。

4．邦塞推广了这个结论。同样的方法可以证明，当 $n\geqslant 5$ 时，存在更强的不等式：

$$P_{n+1}<\sqrt[3]{P_1\cdots P_n}。\tag{9}$$

这个证明的关键在于，在第 2 节末尾讨论的 i 值，对于 n 的三个值实际上是相同的，并且不象刚才证明中的只是倒退二步。

28. 一个改进的不等式

在第 27 篇中我们提到，邦塞的(8)式的证明实际上给出了更好的不等式(9)。事实上，只要增加一个简单的设想，就可以使我们证明一个在某一方面有较多收获的改进的不等式，虽然我们将看到，在其它方面会失去一些东西。

新的设想是这样的：如果M是形如 $6m-1$ 的数，（6 的倍数减 1），那么在M的素因子分解式中，则必然至少出现一个同样形式的素数 $6x-1$。为要证明这是对的，我们注意到，任意一个数必是形如 $6x$，$6x-1$，$6x-2$，$6x-3$，$6x-4$，或 $6x-5$ 这些数之一。因为 $6x$，$6x-2$，或 $6x-4$ 都是偶数，所以 2 是它们之中唯一可以是素数的数。同样，$6x-3$ 可被 3 整除，所以 3 是这类数中的仅有的素数。全部数还剩 $6x-1$ 和 $6x-5$ 这二类，所以所有素数或者是这二种形式中的一个，或是 2，或是 3。然而，2 或 3 不能整除形如 $6m-1$ 的数M。而且，二个数 $6y-5$ 和 $6z-5$的乘积是

$$\begin{aligned}(6y-5)(6z-5)&=36yz-30y-30z+25\\&=6(6yz-5y-5z+5)-5\end{aligned}$$

这仍是原来的形式，因此，若M具有 $6m-1$ 的形式，它的素因子分解式，必须至少含有一个形如 $6x-1$ 的素因子。

为了得到在第27篇基础上改进的新的不等式，我们不象以前那样利用所有的素数$P_1, P_2, P_3, P_4, \cdots = 2, 3, 5, 7\cdots$，而是只取$q_1, q_2, q_3, q_4, \cdots = 2, 3, 5, 11\cdots$。其中$q_1 = 2$，$q_2 = 3$，而其余的$q_s$是形如$6x-1$的素数。

现在我们象以前一样建立表示式$M_1, M_2, \cdots, M_{q_i}$，但我们用$q_s$代替$P_s$：

$$M_1 = q_1 q_2 \cdots q_{i-1} \cdot 1 - 1$$

$$M_2 = q_1 q_2 \cdots q_{i-1} \cdot 2 - 1$$

$$M_3 = q_1 q_2 \cdots q_{i-1} \cdot 3 - 1$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$M_{q_i} = q_1 q_2 \cdots q_{i-1} \cdot q_i - 1。$$

因为我们只在P的位置换写成q，所以有关原来表示式的命题(a)，(b)，(c)仍是对的。

下一步需要稍作解释。恰如以前一样，如果

$$n - i + 1 < q_i, \tag{2*}$$

那么存在某个M，叫它为M_k，不能被素数$q_1, \cdots, q_n$整除。如果M_k本身是素数，它就是形如$6x-1$的素数，并在所有的$q_1, q_2 \cdots q_n$的后面。如果M_k不是素数，根据我们最初的说法，则至少存在一个形如$6x-1$的素数可以整除它。由于这个素数不能是$q_1, q_2, \cdots, q_n$中的任一个，它位于q_n的后面。总之，在任一种情形下，都存在一个形如$6x-1$的可以整除M_k并且是在q_n后面的素数q。由于q_{n+1}是紧接q_n后的一个形如$6x-1$的素数，我们有$q_{n+1} \leqslant q$，并且由于q可以整除M_k，则$q \leqslant M_k < q_1 \cdot q_2 \cdots q_i$。结合这些结果，最后，如果(2*)成立的话，有

$$q_{n+1} < q_1 q_2 \cdots q_i \tag{3*}$$

现在，如我们在第 2 节所作，取可以满足(2*)的 i 的可能的最小值，并列出一个表，能看清对应每一个 n 的 q_i 的值。因为表相当长，我们只写出一部分来：

$n=6$：2,3,**5**,11,17,23

$n=7$：2,3,5,**11**,17,23,29

$n=8$：2,3,5,**11**,17,23,29,···

$n=9$：2,3,5,**11**,17,23,29,···

$n=10$：2,3,5,**11**,17,23,29,···

$n=11$：2,3,5,**11**,17,23,29,···

$n=12$：2,3,5,**11**,17,23,29,···

$n=13$：2,3,5,**11**,17,23,29,···

$n=14$：2,3,5,11,**17**,23,29,···

······················

$n=20$：2,3,5,11,**17**,23,29,···

$n=21$：2,3,5,11,17,**23**,29,···

······················

$n=114$：2,3,5,11,17,23,29,41,47,53,59,71,83,89,**101**,107,113,···

$n=115$：2,3,5,11,17,23,29,41,47,53,59,71,83,89,101,**107**,113,···

··································

$n=121$：2,3,5,11,17,23,29,41,47,53,59,71,83,89,101,**107**,113,···

$n=122$：2,3,5,11,17,23,29,41,47,53,59,71,83,89,101,107,**113**,···

··································

关于这个表，首先要注意的是，无论什么时候 q_i 增加了，它至少增加6。这就允许 i 对于 n 的随后的至少7个值保持同样的值。由于 i 的较长的"停驻"，我们可以证明

$$(q_1q_2\cdots q_i)^6<q_{i+1}\cdots q_n. \qquad (4^*)$$

事实上，如果这个表到某一个 n 的值时是对的，并且如果对于下一个值 $n+1$，i 没有变化，那么(4^*)对于 $n+1$ 仍然成立，因为只是右端增加了。然而，如果 i 变化了。我们可以倒退7步到 $n-6$，使 i 没有变化，并且仍有

$$(q_1q_2\cdots q_i)^6<q_{i+1}\cdots q_{n-6}.$$

如果左端乘以 q_{i+1}^7，右端乘以较大的 $q_{n-5}q_{n-4}q_{n-3}q_{n-2}q_{n-1}q_n q_{n+1}$，我们得到

$$(q_1q_2\cdots q_i)^6\cdot q_{i+1}^7<q_{i+1}\cdots q_{n+1}.$$

二端同除以 q_{i+1}，就得到对于 $n+1$ 的不等式

$$(q_1q_2\cdots q_{i+1})^6<q_{i+2}\cdots q_{n+1}.$$

直到现在，我们还没有完全证明(4^*)。我们只是看到，当 n 到 $n+1$ 时如果它对于 n 是对的，并且 i 没有变化，或对于 $n-6$ 是对的，并且 i 变化的话，那么它对于 $n+1$ 是对的。我们将看到它对于 $n=114$ 和 115 是对的。因为 $n=115$ 是至少向前7步 i 不变化的第一个数，那么(4^*)对于一切 $n\geqslant114$ 都是对的。现在很容易看到(4^*)对于 $n=114$ 和 115 是对的。对于 $n=115$，左端是 $i=16$ 个因子的6次幂，也就是左端是 $6\cdot16=96$ 个因子的乘积，而右端是$n-i=115-16=99$ 个因子的乘积。并且右端的每一个因子大于左边的每一个因子。对于 $n=114$ 的证明是类似的。这个证明是很粗糙的，因为它没有充分利用左端的数比右端的数小得很多这个条件。通过比较仔细的计算，我们或许可以缩小 n 的范

围，但这种计算将是很冗长的。一个非常明显的事实是：(4^*)对于 $n=6$ 是不对的。

如果我们现在在(4^*)的二端都乘以 $q_1q_2\cdots q_i$，并且取 7 次方根，我们有

$$q_1q_2\cdots q_i<\sqrt[7]{q_1q_2\cdots q_n}, \qquad (8^*)$$

如果 $n\geqslant 114$。此式再结合(3^*)，有

$$q_{n+1}<\sqrt[7]{q_1q_2\cdots q_n}, \qquad n\geqslant 114。(9^*)$$

这个式子可以和第 27 篇中的邦塞的不等式(1)及(9)作比较，不过这里是特殊的素数 $q_1,q_2,\cdots$。

为了得到一个象(9^*)那样的对于所有的素数 $P_1,P_2,\cdots$ 都成立的不等式，我们首先注意到，(9^*)的证明表明，特殊的素数 q 序列是无界的。因而，如果 P_r 是任意素数，我们可以求一个 q_n，使 $q_n\leqslant P_r<q_{n+1}$。由于 P_{r+1}是在 P_r 后面的任意形式的第一个素数，所以 $P_{r+1}\leqslant q_{n+1}$。同样地 $q_1q_2\cdots q_n\leqslant P_1P_2\cdots P_r$，因为右端包括所有的素数 $q_1,q_2,\cdots,q_n$ 以及不是这种形式的其它因子。结合这些事实和(9^*)，我们有

$$P_{r+1}<\sqrt[7]{P_1P_2\cdots P_r}。\qquad (9^{**})$$

我们只是对 $P_r\geqslant q_n$，$n\geqslant 114$，也就是对 $P_r\geqslant q_{114}$ 作了证明。这就是我们对邦塞不等式的改进。我们可以说，当 r 充分大时，(9^{**})是对的。通过检验一个素数表，我们可以求得 q_{114}，然后能说明究竟 r 必须多大。甚至(9^{**})对于许多较小的 r 值也许仍然是对的，而这是可以通过实际的计算确定的。但是我们对(9^{**})的主要兴趣在于：如果它对于 r 的某个确定值开始有效，那么继续往后总是对的。至于这个 r 的确切值是什么，则并无多大兴趣。

在(9^{**})的证明中，用显然是较小的七次方根代替了平

方根和立方根，我们改进了邦塞的不等式。然而，在这个过程中，我们也丧失了一些东西。邦塞的不等式除了 n 的最小值外，对于 $n \geqslant 4$ 和 $n \geqslant 5$ 都是对的，而我们证明的(9^{**})只是对于较大的 r 值才成立。

读者可以看到，以上的讨论除了第一篇中已用到的极简单的基础知识外，没有用到任何其它数学知识。证明只依靠纯粹的推理。正因为如此，它清楚地显示了数学是何等困难而巧妙。在某些情形下，需要联系和展开许多分支，才能达到它的目的，但是，在这个只用了很少的一些数学知识，并使推理论证得到发展的例子中，也能揭露出问题的精神实质。如果说，这最后一篇需要一连串困难的思考，表明即使在一窄小的基础上，数学也可以建立起实际而有意义的结构，那么，这也许正好展示了本书的真正动机。

附 录

注释

第 一 篇

欧几里德的说法是(«原本») IX,20,"素数要比任意指定的许多素数还要大。"

其它一些含有无限多个素数的序列，例如，都有公差4的1,5,9,13,···,$4n+1$,···以及3,7,11,15,···$4n-1$,···，可以用初等方法证明。然而，有相同公差4的序列2,6,10,···，$4n+2$,···，由于它的所有项都是偶的，所以只含有单独的素数2。一般说来，有任意公差的序列，如果第一项与公差互素，那么定理成立。这已在迪里茨拉特的一篇著名而艰深的论文中，借助较高等的数学得到证明。

第 二 篇

曲线网络若表示公共汽车线路，则曲线网络是在一个平面内。但这不是必须的。本节的全部结果对于空间的曲线网络也是成立的。但对照起来和本篇很类似的第10篇的论题，其中的结论不能毫无改变地适用于空间。

第 三 篇

第2节。这个讨论可以由圆直接转移到椭圆。等边三角

形的作用被这样的三角形所取代，即过该三角形的每个顶点的切线平行于对边。

第3节。这个预备步骤是由斯坦尼茨提出的。

第 四 篇

这第一个证明在希腊数学家的著作中是找不到的，但这是他们已能作出并已经知道的典型的证明。第二个证明见欧几里德的，«原本»X，但在亚里斯多德的著作中已有记载，似乎说明它比欧几里德更早。

第五篇和第六篇

施瓦茨和斯坦纳的著作中的判断是对平面和球面三角形同时给出的。

三角形是锐角三角形显然是必要的，因为只有在这种情形下，所有的高线在三角形内部。在钝角三角形内，垂足三角形不是严格的内接三角形。在直角三角形内，垂足三角形退化为直线。

在施瓦茨的证明中，锐角三角形这个事实，在垂足三角形内接于原三角形的假设中用到。在费瑞的证明中，直到最后垂足三角形也没有出现。三角形的锐角性在下述方面起着作用，即保持角 $U'AU''$ 比二个直角小，使 $U'U''$ 和 AC，以及和 AB 的交点M和N在这个边上，而不在它们的延长线上。并且由于三角形是锐角，垂足E在 BC 边上，而不在它的延长线上。

费瑞证明的优点，是它在非欧几何内也是成立的。而施瓦茨证明则是不对的，因为它利用了三角形的三内角和是180°的事实，以及欧几里得的平行线的概念。特别是，因为锐角球面三角形的边小于$\frac{\pi}{2}$，所以费瑞的证明对于球面三角形的每步仍然成立。

第 八 篇

维戈·勃鲁恩利用了比组合数更一般的组合函数。

第 九 篇

巴夏特(Bachet)在编辑《数学》一书时，声称《数学》的某一部分隐含了一个数表示成四个平方数之和的定理，费马在一封信里给出了证明的梗概，而欧拉和拉格朗日曾证明了本篇中叙述的定理。

韦林猜想9个立方数，19个四次幂，等等是充分的。

韦林的猜想中的"等等"可以解释成这样的意思，任何自然数n可以表示成K次幂的和的最大个数为

$$I=2^K+q-2$$

其中q是不超过$\left(\frac{3}{2}\right)^K$的最大整数。那么，对于数

$$n=2^Kq-1$$

确实需要好多个K次幂，在此，因为它小于3^K，可以利用

的仅被加数 1^K 和 2^K。现在迪克逊和比拉在 1936 年及下一年证明了，如果（用稍后的尼文的改进式）

$$\left(\frac{3}{2}\right)^K - q < 1 - \frac{q}{2^K},$$

对于 $K \geqslant 6$，K 次幂的数 I 对于所有的 n 确实是充分的。这个条件对于 $2 \leqslant K \leqslant 400$ 成立。是否它对所有的 K 都成立，我们还不知道。然而对于可以不成立的 K 次，需要 K 幂的数 $g(K)$ 已被确定。对于 $K = 2,3$，数 I（分别 $=4,9$）还是好的。但对于 4,5，我们现在仅知道

$$19 \leqslant g(4) \leqslant 35,\quad 37 \leqslant g(5) \leqslant 54。$$

迪克逊和比拉的工作完全停留在维诺格拉道夫的结果上，这个结果是用他的由哈代和里德尔伍特的方法发展来的函数论方法获得的。

第 十 篇

第 4 节中，没有二重点的闭曲线，把平面分成二个区域不是自明的。这个定理对一切曲面是不对的。例如，在轮胎形曲面上就是错误的。因此这个定理必然被说成是在平面上是正确的，而且还需要证明。有一个首先由乔旦给出的证明，这个定理就以他的名字命名。由于乔旦的定理对轮胎不能成立（图 117），我们的全部论证对这个曲面也就是错误的。事实上，曲线在轮胎上可以使其二重点的次序是 1212（图 118）。

图 34 的结可以不撕开它而变形为图 33 的结，即一个交错结。但是认为任何结都能变形为一个交错结是不正确的。



图 117



图 118

彭克威茨曾给出了一个其投影不是交错结的例子。

第十一篇

欧几里德没有明显指出素数分解的唯一性定理。但是他证明了，如果乘积可被素数整除，那么必有一个因子能被此素数整除(«原本»Ⅶ，24，29)，由此素数分解的唯一性定理随之得出。在此证明中，欧几里德利用了最大公约数以代替最小公倍数。

二十世纪分别由舍曼罗，汉斯，丘威尔找到的一个利用数学归纳法的简单证明，将在这里指出。

开头几个自然数一定有唯一的素因子分解。其实，除去作为单位的 1，数 2 和 3 都是素数。假设存在着 有二个不同素数分解的数，那么它们当中必有最小的一个，设为

$$N=P_1\cdot P_2\cdots P_K=q_1\cdot q_2\cdots q_L$$

其中 P 和 q 是素数，我们假定这每一个乘积 都 是 按数值大小次序排列

$$P_1\leqslant P_2\leqslant\cdots\leqslant P_K \text{ 和 } q_1\leqslant q_2\leqslant\cdots\leqslant q_L.$$

所有的 P 都和所有的 q 不同， 因为如果任意一个 $P_i=q_j$，那

么可以消去这个因子，从而得到$\frac{N}{P_1}$，这是比N小且有二个不同的素因子分解的数。不失一般性，我们可以假设$P_1<q_1$。然后组成

$$M=P_1\cdot q_2\cdots q_t<q_1\cdot q_2\cdots q_t=N.$$

这个M和N一样可被P_1整除。因而差$N^*=N-M$也可被P_1整除，这样，就有了一个含有P_1的素数分解。另一方面，我们有

$$N^*=(q_1-P_1)q_2\cdots q_t,$$

q_1-P_1不能由P_1整除，而且没有一个素数$q_2,\cdots,q_t$等于P_1。这就意味着N^*还有一个不含P_1的素因子分解。这样，就有了二个不同的素因子分解。但$N^*<N$，与N是这类数中的最小者的假设矛盾。这个矛盾说明，存在具有二个不同的素因子分解的数是错误的。

第十三篇

第10节中几何图形在被搞乱（但未被撕破）的情况下仍保持不变的性质，形成数学中一个分支的主题，即拓扑或形势几何。我们研究的正多面体问题是纯拓扑问题。

把多面体充气为球的概念，表示的是一个特殊的拓扑假定。如果研究拓扑等价于轮胎的多边形，我们将得到完全不同的结果，并且会显示出第7节到第10节的拓扑意义。已经证明，在轮胎形曲面上，存在无限多个“正”地图（在拓扑意义下），但象在度量几何意义下的正多边形，这是不能实现的。

在轮胎形曲面上，欧拉公式成为（见12篇，第5节）

(3*) $$v-e+f=0。$$

方程(4)和(5)在这里显然成立，所以(6)变成

(6*) $$f(2\varphi+2\varepsilon-\varphi\varepsilon)=0。$$

并且，因为$f\neq 0$，则

(7*) $$2\varphi+2\varepsilon-\varphi\varepsilon=0,$$

或

(9*) $$(\varphi-2)(\varepsilon-2)=4.$$

代替不等式(9)，现在是一个等式。4分解成为二个因子的乘积只可能是$1\cdot4$，$2\cdot2$，和$4\cdot1$，所以$(\varphi-2)\cdot(\varepsilon-2)$也仅可能有三种情形。这样，对于$\varphi$和$\varepsilon$可得到表

φ	3	4	6
ε	6	4	3

这些值满足(9*)，因而也满足(7*)。那么f的任意数值都满足(6*)，因为对于一切f，都有$f\cdot0=0$。所以，在这种情形，f, v和e不能由φ和ε的值求得。

其实，对于每一对φ, ε，有f, v和e的无限多个值。

对于一对$\varphi=4$，$\varepsilon=4$，选取任意一个数$a>1$，并且把小正方形排列成一个较大的正方形（图119），并可卷成一个圆筒（图120），而圆筒可弯成轮胎（图121）。那么这轮胎形曲面由“正”地图所覆盖，在这个“正”地图上，每个国家有$\varphi=4$条边界，并且在每个顶点有$\varepsilon=4$个国家相交。这里我们有$f=a^2$，并且很容易看到，也有$v=a^2$，$e=2a^2$。因为



图 119

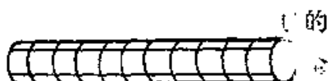


图 120

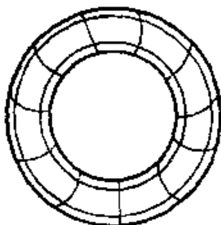


图 121

a 是大于 1 的任意数，所以在轮胎形曲面上，有无限多个拓扑"正"地图。(对 $\varphi=6$，$\varepsilon=3$ 构作地图是比较困难的，$f=7$，$v=14$，$e=21$ 的情形会出现在这些地图中。这是第 12 篇第 5 节中提到过的地图，与轮胎上的着色问题有关。)

这些正多面体不能在度量几何意义下实现。因为如果 $\varphi=4$，$\varepsilon=4$，那么面必须是正方形。但四个正方形围绕一点放在一起，只可能是平面，并且不会形成三维顶。许多正方形不管是怎样放在一起的，它们将连续铺在一个平面内，并且不会形成一个立体的曲面。对 $\varphi=6$，$\varepsilon=3$ 同样是对的，因为是三个正六边形围绕一点平放在一个平面内。类似，对 $\varphi=3$，$\varepsilon=6$，6 个等边三角形围绕一点平放在一平面内。

扼要地说，我们注意到，首先，在球形曲面上只存在 5 种拓扑正多边形，是球的拓扑性质。轮胎形曲面上存在无限多个拓扑正多边形。第二，拓扑的规律不能推导出度量的规

律。如果5种度量性的多边形在球面上可以实现，那是由于特殊的度量性质所致。

第十四篇

除第7节 $n=4$ 的情形（已在第8节中讨论），已相当充分地证明了 $x^n+y^n=z^n$ 对整数且同时是素数的 $n=P$ 的不可能性。在有关文献中，区分了二种情形：情形Ⅰ，$x\cdot y\cdot z$ 不能被 P 整除，情形Ⅱ，数 x，y，z 中的一个（并且只有一个）能被 P 整除。在情形Ⅰ中费马猜想已对所有素数 $P<253,747,889$ 得到证明。在情形Ⅱ，用到较深的数论中的定理，并且里莫尔和伊玛·里莫尔和凡迪伐尔利用在洛杉矶的一种电子计算机，验证了费马猜想对所有 $P\leqslant 4001$ 成立。

第十六篇

荣格研究了 n 维中的类似问题。他把

$$r=d\sqrt{\frac{n}{2(n+1)}}$$

当作 n 维有限点集跨度球的半径，其中跨度为 d。关于 $n=2$ 的情形，就是本文所提的问题

第十九篇

第7节，若 P 是素数，则形为 2^P-1 的素数叫作默生尼素数。现在知道的这些数有

$P=2,3,5,7,13,17,19,31,61,89,107,127,521,607,1279,$ $2203,2281$，其中最后 5 个，在 1952—1953 年借助洛杉矶数字电子计算机已经求得。每一个默生尼素数产生一个完全数。

第二十二篇

本篇末尾指出的证明已由卡拉萨杜雷斯达台完成。一个不同的完整证明是埃特尔提出的。

第二十三篇

高斯讨论了循环小数，但他利用了数论中的很多结果。

第 4 节。数 $\varphi(n)$ 由 n 而定。它通常叫作欧拉函数，并且可以定义为分母是 n 的既约真分数的个数。

第 8 节。如果 $P, q, r, \cdots$ 是不同的可整除 n 的素数，那么 $\varphi(n)$ 由公式

$$\varphi(n)=n\cdot\frac{P-1}{P}\cdot\frac{q-1}{q}\cdot\frac{r-1}{r}\cdots$$

给出。

这个公式的证明，可在任一本教科书中找到。

第二十五篇

第 7 节中等宽度曲线的性质，由定理Ⅳ可断定对一切凸曲线也能成立，并且也可以直接证明。但是这个证明要利用

极限过程，并且最好在它的解析公式中完成。

等宽度曲线可以由初等方法处理的理论部分，本篇中已全部给出。

第二十六篇

第 1 节：马歇罗尼在他的著作中说，由于实用的缘故，他已研究了只用圆规的作图问题。只用圆规作图，常比用直尺作图精确，事实上，如马歇罗尼知道的，这就是天文学家使用他们的仪器得到所希望的精确度时的情形。他由于研究圆规作图，导致他发现了所有欧几里德的作图，都可以只用圆规完成的结果。

马歇罗尼把他的著作送给了拿破仑，赞扬他是北意大利的解放者。拿破仑转而带着这本书与法国科学院的院士们作了谈话（1797 年，12 月 10 日），引起法国学者的注意。

斯坦纳，生于瑞士，是柏林大学的教授。在他的著作中已讲到本篇所提到的问题。

第 6 节和第 7 节。我们考察的只是二个非相交圆的情形。找一个对于二个相交圆的我们用过的那类投影是不可能的。事实上，这样的一个投影，应证明圆心不可能由只用直尺的作图求得。但在这种情形下，如下的作图是可以求得的。

这个作图实际很简单。在第一个位置，图 122 说明如果给出二个平行弦 AA' 和 CC'，圆的直径可以怎样找到。由于对称，直线 DD' 通过圆心。如果第二对平行弦也已经知道，那么第二个直径也能画出，并且圆心由二条直径的交点

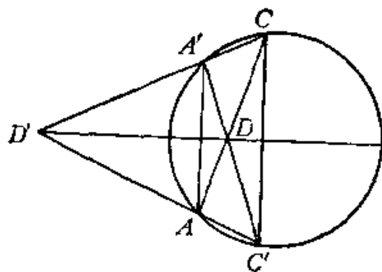


图 122

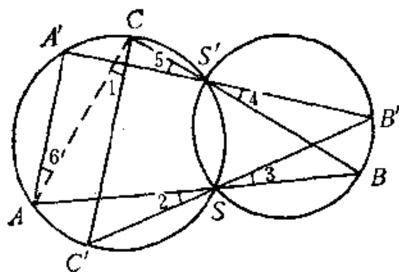


图 123

决定。现在必须求作二条平行弦的方法。这如图 123 所示。直线 AA' 是任意一条与圆相交的弦。开始是在 A，我们画直线 ASB 和直线 $BS'C$，这样确定交点 C。类似，开始在 A' 点，我们画 $A'S'B'$ 和 $B'SC'$。那么点 C 和 C' 决定另一条弦。要证明 AA' 和 CC' 是平行的，只需证明内错角 1 和 6 是相等的。逐个考察角 $1,2,\cdots,6$，可见每一个都与下一个相等，这或者是因为它们对同一弧或者因为是对顶角。

三个非相交圆确定圆心的证明稍微有点复杂。我们在这里将不转载，因为这需要更多的几何知识。

第 8 节。另外，可以找到定理的一个较短的证明，但和我们已经给出的证明一样，都不具有纯几何的样式。我们所以选取这个证明，是因为它强调了定理的实质部分，即斜圆锥的对称性。

第二十七篇

实际上现在已经知道更多的 $P_{n+1}<2P_n$。然而，即使是最好的结果，仍不足以肯定是否在每一对依次相邻的平方数之间总存在一个素数，例如，在 100 和 121 之间，121 和 144 之间，如此等等。